
Usability-Testing mobiler Szenarien als Sekundärtask – Vergleich technischer Ansätze

Kerstin Blumenstein & Grischa Schmiedl

FH St. Pölten, Institut für Creative\Media/Technologies

Zusammenfassung

Wird eine Applikation auf einem Smartphone nebenher benutzt, so muss der User seine kognitiven Ressourcen auf das Benutzen der Applikation und das Ausführen einer primären Tätigkeit aufteilen. Labor-basierende klassische Usability-Tests berücksichtigen nicht den Einfluss des Primärtasks. In Feldstudien können jedoch die den Test beeinflussenden Faktoren nur schwer gemessen und kaum gesteuert werden.

In (Schmiedl u. a., 2011) stellten wir eine Testmethode vor, welche die Anforderungen erfüllt. Mit einem Autorennspiel wird ein steuerbarer Primärtask generiert und gleichzeitig vom User die zu testende mobile Applikation bedient. Anschließend werden die Dauer und Anzahl der Interaktionseinheiten gemessen.

Um zu weiteren Erkenntnissen über die Testmethode zu gelangen, vergleichen wir in diesem Paper zwei technische Ansätze zur Überwachung des Szenarios. Ein Low-Budget-Setup (Kopfkamera) wird mit einem High-Budget-Setup (mobiles Eyetracking) verglichen. Beide Setups eignen sich, um unser gewähltes Szenario zu überwachen und entsprechende Daten auszuwerten. Jedoch ist zu erkennen, dass bei der von uns getesteten Technik das High-Budget-Setup keinen Mehrwert gegenüber den Low-Budget-Setup liefert.

1 Einleitung

Das Smartphone ist zu einem uns ständig begleitenden Medium geworden. Es ermöglicht, die stetig steigenden Anforderungen der Gesellschaft nach Information zu befriedigen. „Immer informiert“ ist nach „ständig erreichbar“ ein kritischer Erfolgsfaktor vieler Knowledge-Worker geworden. Die Szenarien, in denen Applikationen genutzt werden, haben sich dadurch verändert.

Ein modernes Smartphone, ausgestattet mit GPS-Navigation, Kamera und Highspeed-Internetzugang, ist ideal geeignet, um auch in Bewegung genutzt zu werden. Manche Applikationstypen (z.B. POI-Applikationen, Navigationssoftware...) sind sogar explizit für mobile Szenarien vorgesehen. Der User kann sich in solchen Nutzungsszenarien jedoch nicht nur auf die Benutzung der Applikation am Smartphone konzentrieren, sondern muss auch auf die Umgebung achten. Die Interaktion mit dem Gerät ist in kurze Einheiten geteilt. Die Nutzungsszenarien eines Smartphones unterscheiden sich somit von denen anderer Medien, wie bspw. einem Desktop-Rechner oder einem Laptop.

Demnach sollen beim Testen Usability-Methoden angewendet werden, die speziell für das Medium „Smartphone“ angepasst wurden. Usability-Testing ist ein wichtiger Teil des Evaluierungsprozesses. Ein valider Test muss folgende Anforderungen erfüllen:

- Der Test soll in einem realistischen Szenario stattfinden.
- Der Test selbst und alle relevanten Parameter müssen überwacht werden.
- Der Test und dessen Resultate müssen reproduzierbar sein.

Bei Usability-Tests mobiler Applikationen entstehen dadurch mehrere Probleme. Das Testen mobiler Applikationen im Labor ist kein realistisches Szenario. Feldtests sind jedoch schwer zu überwachen. Die Umfeld-Parameter sind nicht immer beeinfluss- und reproduzierbar.

2 Bestehende Forschungsansätze

Bereits 1999 fanden Kristoffersen und Ljungberg heraus, dass die technischen Beschränkungen (kleiner Bildschirm, langsame Internetverbindung...) nicht das alleinige Problem von mobilen Geräten ist. Das mobile Gerät teilt sich bei der Nutzung die physischen Ressourcen sowie die Aufmerksamkeit des Users

mit einem anderen Task. Die Konsequenz daraus ist: der User muss seine Arbeitssituation auf das mobile Gerät anpassen. Ein Autofahrer bspw. muss seinen Wagen stoppen, um eine eMail zu lesen und zu beantworten. (Kristoffersen & Ljungberg, 1999)

Bei einer Forschungsreihe im Jahr 2000 von Kim et al. wurden 37 Testpersonen mit einem internetfähigen Mobiltelefon ausgestattet. Sie bekamen die Aufgabe, es zu nutzen, wann immer sie wollen. 1.552 Sessions wurden anschließend analysiert, um herauszufinden, in welchem Kontext die Testpersonen das mobile Internet tatsächlich verwendeten. Kim et al. dokumentierten, dass das mobile Internet hauptsächlich in erfreulichen Situationen sowie in stiller und ruhiger Umgebung genutzt wird. (H. Kim u. a., 2002) Das Resultat dieser Studie zeigt, dass kein Bedarf bestand, Applikationen in Szenarien mit geteilter Aufmerksamkeit zu testen. User benutzten mobile Applikationen nicht nebenher.

Fünf Jahre später berichteten Zhang & Adipat, dass das Testen von Usability mobiler Applikationen im Labor keine überzeugenden Ergebnisse liefert. Experimente in Feldtests erlauben jedoch kaum Kontrolle über die Testpersonen und die Testumgebung und somit keine verlässlichen Resultate. Zhang und Adipat schlussfolgerten, Usability-Testmethoden, die sich für Desktop-Applikationen bewährt haben, können nicht 1:1 für das Testen von mobilen Applikationen übernommen werden. Die Frage nach dem Wie wird im Paper offen gelassen. (Zhang & Adipat, 2005)

2007 testeten Looije et al. mobile Karten-Applikationen. Sie stellten fest, dass es bei der Nutzung der Applikationen mit mobilen Geräten nötig ist, diese zu testen, während der User in Bewegung ist. Dies macht jedoch die Evaluation schwierig. Looije et al. stimmen mit Zhang überein, dass eine Methode für realistisches Usability-Testing, welche die Probleme von Feldstudien vermeidet, wünschenswert ist. Ihr Fazit lautete: Es existieren noch viele Herausforderungen für Lösungskonzepte, sowohl in Bezug auf Design, als auch auf Usability-Testing für mobile Geräte. (Looije u. a., 2007)

Coursaris & D. Kim erarbeiteten 2007 eine Forschungsagenda für mobile Usability. Sie durchforsteten publizierte Artikel und Paper, die Usability mobiler Applikationen testeten, um zu eruieren, wie viele Studien welche Testarten nutzen. Das Ergebnis: 58% aller Studien finden im Labor statt. Nur 22% nutzen die Methode der Feldstudie. 11% benutzen beide Methoden, um bessere Resultate zu gewinnen. In 9% der Studien war die Methode nicht dokumentiert. (Coursaris & D. Kim, 2007)

Perry & Hourcade beschrieben 2008 ihre Bemühungen, um die Usability für mobile Applikationen auf Touchscreen-Geräten zu verbessern. Nicht alle Positionen auf dem Bildschirm sind für Menschen gleich gut zu erreichen. Dies ist durch die physischen Grenzen der Hand bedingt. Im Test sollten die Testpersonen vordefinierte Positionen auf dem Display berühren. Dabei wurden die Genauigkeit und die Geschwindigkeit gemessen. Die Ergebnisgrafik zeigt bevorzugte und weniger bevorzugte Regionen, die für Interaktionselemente genutzt werden sollten. Die Autoren mussten allerdings während ihrer Studie feststellen, dass sie nicht den Kontext berücksichtigten, in dem die Geräte tatsächlich genutzt werden. Sie vermuteten, dass der User auf einer belebten Straße möglicherweise nicht so gut abschneiden würde, wie im Labortest. (Perry & Hourcade, 2008)

Schließlich versuchten Oulasvirta et al. mehr über reale mobile Szenarien herauszufinden. Eines ihrer wichtigsten Ergebnisse war, dass Applikationen, die in Szenarien mit geteilter Aufmerksamkeit genutzt werden, die Zeit beschränken sollte, die ein User für eine Interaktionseinheit benötigt. (Oulasvirta u. a., 2005) In ihrer Arbeit analysierten sie unterschiedliche Situationen mit dem mobilen Gerät. Sie erkannten, dass auch in Situationen wie Sitzen in einem öffentlichen Bus, die durchschnittliche Länge der Interaktionseinheit für das mobile Gerät nur sechs Sekunden beträgt.

Verschiedene Arbeiten beschäftigen sich mit dem Finden eines geeigneten Equipments für Labor und / oder Feldstudien mit mobilen Geräten und Applikationen. Schusteritsch et al. präsentierten die Infrastruktur, die Google Labs für mobile Usability-Tests verwendet. Im Paper konzentrierten sie sich auf verschiedene Möglichkeiten zum Aufnehmen des Bildschirms und Erfassen der Interaktionen. (Schusteritsch u. a., 2007) Eine professionelle Hardware-Konfiguration für mobile Usability-Feldstudien wurde von Oulasvirta und Nyysönen in (Oulasvirta & Nyysönen, 2009) vorgestellt. Das Equipment ist transportabel, in zwei Aluminiumkoffern verpackt und kostet rund 10.000 Euro.

Eine wichtige Erkenntnis muss Kallio und Kaikkonen zuerkannt werden. Sie verglichen Labor- und Feldstudien von mobilen Applikationen und beobachteten, dass in den getesteten Anwendungen tatsächlich die gleichen Usability-Probleme gefunden wurden, egal welche Methode benutzt wird. (Kallio & Kaikkonen, 2005) Die Auswirkung, die ein bestimmtes Problem in einem mobilen Szenario hat, ist dennoch nicht unbedingt die gleiche, wenn die Applikation in einer ruhigen und kontrollierten Laborumgebung getestet wird. Dennoch ist ein Labortest aussagekräftig, er muss allerdings mit zusätzlichen Methoden vervollständigt werden.

In (Schmiedl u. a., 2011) stellten die Autoren eine realitätsnahe Methode zum Testen mobiler Usability vor. Dieses Setting ist eine mögliche Antwort auf Zhang's Frage, wie Usability von mobilen Multimedia-Applikationen tatsächlich evaluiert werden kann. Bei dieser Methode wurde ein handelsüblicher Autorennsimulator verwendet, um den benötigten primären Ablenkungstask zu erzeugen. Die Bedienung der mobilen Applikation wird dabei zum sekundären Task im Sinne der Teilung der kognitiven Ressourcen der Testperson.

3 Forschungsfragen

In diesem Paper setzen wir die Arbeit aus (Schmiedl u. a., 2011) mit einem Vergleich von Low-Budget- und High-Budget-Equipment fort. Für das Low-Budget-System kommt das Equipment zum Einsatz, das bereits im Test für (Schmiedl u. a., 2011) verwendet wurde. Der High-Budget-Vergleich wird mit einem professionellen mobilen Eyetracking-System der Firma SMI durchgeführt.

Unsere Forschung fokussiert auf den Vergleich der beiden technischen Varianten:

- Mit welcher Variante kommt der User besser zurecht? Der Test soll durch das Equipment nicht beeinflusst werden oder den User irritieren.
- Welche Daten sind mess- und nutzbar?
- Können mit Hilfe des deutlich teureren mobilen Eyetracking-Systems zusätzliche bzw. genauere Daten gemessen werden, die den höheren Anschaffungspreis rechtfertigen? (z.B. Wie lange benötigt der User bis der Button gefunden wird?, Wo genau schaut er hin?, ...)
- Welche technischen Probleme treten bei der Anwendung im mobilen Szenario auf?

Es erfolgt eine Aufwandsmessung im wirtschaftlichen Sinn. Relevant sind für uns die Kosten der Lösung, die Zeit für den Aufbau und die Unsicherheitsfaktoren.

4 Hardware Setup

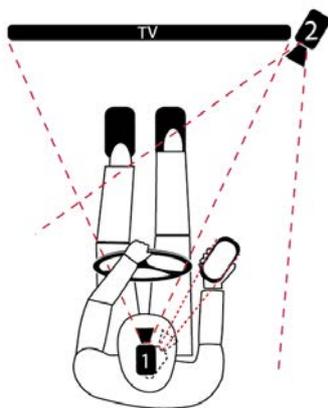


Abbildung 1: Hardware Setup

Der verwendete Simulator setzt sich aus einer normalen Spielekonsole (Sony Playstation 3), einem Autorennspiel (Grand Turismo 5) und einem 47" LED-Fernseher zusammen. Das Spiel unterstützt einen Übungsmodus, der es ermöglicht, allein auf der Fahrbahn zu sein. Um das Fahrgefühl zu stärken, haben wir ein Lenkrad sowie Brems- und Gaspedale (Logitech Driving Force GT) angeschlossen und an einem justierbaren Lenkradtisch (Speedblack EVO) fixiert.

Der Aufbau des Grundsettings ist aus Abbildung 1 ersichtlich. Der Test wird mit zwei Kameras aufgezeichnet. Kamera 2 befindet sich direkt neben dem Fernseher und ist frontal auf die Testperson gerichtet. Es handelt sich um einen handelsüblichen Camcorder. Kamera 1 unterscheidet sich in den beiden Testsettings.

Um Reflexionen auf dem Mobiltelefon zu vermeiden, entscheiden wir uns für indirekte Beleuchtung.

Um Missverständnisse zu vermeiden, muss klargestellt werden, dass als primärer Task jede beliebige Aufgabe verwendet werden kann, solange diese sitzend oder zumindest relativ stationär durchgeführt

werden kann - da sonst die Beobachtung erschwert wird. Außerdem muss der Primärtask einerseits einen hohen Ablenkungsfaktor aufweisen, andererseits aber dennoch weitgehend automatisiert von der Testperson – einhändig – durchführbar sein. Ein Autorennspiel ist eine kostengünstige Variante zur Erzeugung dieses Ablenkungstasks im Simulator, die alle Kriterien erfüllt. Die tatsächliche und in vielen Ländern illegale Nutzung eines Handys während des Autofahrens soll damit aber weder simuliert noch empfohlen werden.

4.1 Low-Budget-Equipment

Für das Low-Budget-Setting nutzen wir Equipment, das bereits in (Schmiedl u. a., 2011) verwendet wurde. Es handelt sich um das Modell GoPro HD, eine robuste, weitwinklige Kopfkamera, die vor allem im Sport- und Outdoor-Bereich seine Anwendung findet. Mit einem entsprechenden Kabel wird diese mit einem Fernseher verbunden, um das Bild während des Testens überwachen zu können. Die Aufnahme für die spätere Analyse erfolgt direkt von der Kamera auf eine SD-Karte.



Abbildung 2: Low-Budget-Testsetting

4.2 High-Budget-Equipment

Als High-Budget-Equipment kommt das mobile Eyetracking HED Vision der Firma SMI Vision zum Einsatz. Das Eyetracking-System wurde auf eine Kappe montiert. Das Eyetracking-System eröffnet mittels zweier Kameras die Möglichkeit, die Augenbewegung der Testperson sowie die Umgebung in Blickrichtung aufzuzeichnen und damit die vom Benutzer fokussierten Bildbereiche zu errechnen. Das System ist via USB-Kabel mit einem Laptop verbunden. Auf diesem kommt für die Hardwaresteuerung und für die Aufzeichnung die Software „iViewX HED“ zum Einsatz. Die Kamera zur Umgebungsaufnahme ist mit Wechselobjektiven ausgestattet. Unsere Wahl fällt auf eine 104°-Linse. Mit diesem Objektiv ist es möglich sowohl den Fernseher in der Ferne als auch das Handy in der Nähe auf einem Bildausschnitt zu haben.



Abbildung 3: High-Budget-Testsetting

4.3 Kosten

Die Gesamtkosten unterteilen sich in zwei Bereiche:

1. Die Kosten des Basissystems (Simulator und Kamera 2) belaufen sich auf 1.460 Euro. Diese Kosten fallen in beiden Testsettings an.
2. Die Kosten für das jeweilige Testsetting: Für das *Low-Budget-Setting* fallen hier Kosten in Höhe von 535 Euro an. Enthalten sind die Kopfkamera (GoPro HD und SD-Karte) und das notwendige offene Gehäuse (GoPro HD Skeleton Gehäuse) sowie der Überwachungsmonitor.

Beim *High-Budget-Setting* belaufen sich die Kosten auf rund 24.500 Euro. Dabei ist sämtliches Equipment enthalten, das zum Test verwendet wurde.

Die Kostenaufstellung kann Tabelle 1 entnommen werden.

Tabelle 1: Kostenaufstellung¹

SIMULATOR & GESAMTÜBERWACHUNG			
Playstation 3	250 €		
Lenkrad für PS3 & Lenkradtisch	200 €		
Fernseher	800 €		
Camcorder & Stativ	210 €		
Zwischensumme:	1.460 €		
LOW BUDGET		HIGH BUDGET	
GoPro HD	350 €	9.500 €	Mobiles Eyetracking & Kappe
GoPro HD Skeleton Gehäuse	45 €	~15.000 €	Laptop & Software
Fernseher (Überwachung GoPro)	120 €	30 €	Zusätzliches Objektiv
SD-Karte 16 GB	20 €		
	1995 €	25.990 €	

5 Testablauf

Der Test ist in vier Teile gegliedert:

1. Simulator, Gerät und Szenario kennen lernen

In diesem Teil bekommt die Testperson die Möglichkeit, sich mit dem Simulator und dem Testgerät (iPhone 4) vertraut zu machen. Zudem wird die Person in das Testszenario eingewiesen und kann sich mit der Applikation vertraut machen.

2. Referenz-Test

Die Testperson muss das Szenario störungsfrei (ohne Simulator) absolvieren.

3. Simulator-Test: Low-Budget

Die Testperson wiederholt das Szenario aus dem Referenz-Test, während sie im Simulator fährt. Die Überwachung erfolgt durch die am Kopf des Probanden angebrachte Kamera.

4. Simulator-Test: High-Budget

Das Szenario wird erneut wiederholt. Die Überwachung erfolgt mit mobilem Eyetracking.

Da eine Testperson das Szenario drei Mal wiederholen muss, suchten wir ein Szenario, mit dem sich jede Testperson bereits im Vorfeld auskannte. Eine Verbesserung der Leistung im Laufe des Tests durch einen Lerneffekt sollte vermieden werden. Die Wahl fiel auf die Applikation Facebook. Die Testpersonen sollten im eingeloggt Zustand zwei kurze Statusmeldungen posten. Sie mussten die iPhone-Applikation öffnen, eine Statusmeldung posten, anschließend in den Browser wechseln und eine weitere Statusmeldung in der Desktop-Website posten.

6 Durchführung und Ergebnisse

An unserem Test nahmen acht Testpersonen im Alter von 21 bis 40 Jahren aus dem universitären Umfeld teil. Sechs von acht Personen waren Brillenträger. Zwei Personen waren weiblich, sechs männlich.

6.1 Zeiten

Tabelle 2 zeigt die Dauer für Aufbau und Einrichtung der beiden Testsettings. Der Aufbau des *Low-Budget-Setups* betrug im Durchschnitt 1:44 Minute. Das *High-Budget-Szenario* dauerte im Durchschnitt

¹ Als Richtwerte wurden die aktuellen Preise Amazon (<http://www.amazon.at>), RS Components Handelsges.m.b.H. (<http://at.rs-online.com/>), Cam For Pro (<http://www.camforpro.com/>) entnommen. Die Preise sind gerundet.

6:28 Minuten und benötigte somit 4:44 Minuten länger für Aufbau und Einrichtung als das Low-Budget-Szenario. Bei den Personen sieben und acht konnte das High-Budget-Szenario nicht absolviert werden, da das Kalibrieren des mobilen Eyetrackings fehlschlug. Nach 10 Minuten wurde abgebrochen.

Sowohl die Zeiten für das Einrichten des Low-Budget-Setups als auch des High-Budget-Setups schwankten. Beim Low-Budget-Setting ließ sich diese Schwankung mit dem Anpassen der Befestigungsgurte an die Kopfgröße begründen. Beim High-Budget-Setting stach Testperson 2 mit 6:46 Minuten hervor. Diese Person trug eine Mütze. Im Low-Budget-Setting stellte dies kein Problem dar. Im High-Budget-Setup stellten wir fest, dass eine richtige Positionierung der Augenkamera nicht möglich war. Die Mütze musste abgenommen und das Setup neu eingerichtet werden.

Tabelle 2: Dauer (in Minuten) für Aufbau und Einrichtung

	TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8	Durchschnitt
Low-Budget	02:08	01:26	01:12	01:00	01:18	02:28	01:50	02:33	01:44
High-Budget	04:19	06:46	05:11	06:12	03:47	05:05	10:15 A	10:10 A	06:28
Differenz	02:11	05:20	03:59	05:12	02:29	02:37	08:25	07:37	04:44

6.2 User

Drei der acht Testpersonen (37,5 %) bevorzugten das Low-Budget-Testsetting (Kopfkamera), drei weitere (37,5 %) das High-Budget-Testsetting (mobiles Eyetracking). Zwei Personen (25 %) fanden keines von beiden angenehm.

Low-Budget

Das Hauptproblem der Kopfkamera war die Fixierung. Drei Testpersonen beklagten, dass die Kamera auf der Stirn drückt. Dies kann durch schlechtes Fixieren der Kamera am Kopf hervorgerufen werden. Bei kleineren Köpfen muss das Band der Kopfkamera sehr eng zusammengeschnürt werden, die Fixierung kann nicht optimal erfolgen. Das Gewicht von 167g (inkl. Gehäuse) drückt unangenehm auf die Stirn. Eine Testperson beschrieb das Testsetting als unnatürlich, weil sie den gesamten Kopf bewegen musste, damit die Blickbewegungen deutlich erkenn- und auswertbar waren.

High-Budget

Für drei Personen war die mobile Eyetracking-Kappe angenehm zu tragen. Den Spiegel (vor dem linken Auge) und die Kamera (über dem rechten Auge) empfanden drei der acht Testpersonen als störend. Zwei Personen empfanden das Testsetting mit mobilem Eyetracking anstrengender. Mit dem Wissen, dass die Augenbewegungen aufgezeichnet werden, sahen sie nach ihren Angaben bewusster hin.

6.3 Daten & Auswertung

Mit der verfügbaren Softwarelösung „iViewX HED“ der High-Budget-Lösung war eine automatische Target-Erkennung nicht möglich. Daher war es für beide Testsettings notwendig, eine manuelle Messung der Interaktionseinheiten (Anzahl und Dauer der Interaktionseinheiten, Gesamtdauer der Interaktionseinheiten, Gesamtdauer des Szenarios) durchzuführen. Für diese Auswertung wurde die iPhone-Applikation MultiTimer² verwendet. Sie unterstützte bei der Messung der Interaktionseinheiten, während man das Video anschaut. Eine sichere Messung ermöglichte in beiden Testsettings Kamera 2 (vgl. Abbildung 1: Hardware Setup). Hier waren die Blicke der Testpersonen deutlich erkenn- und auswertbar.

Low-Budget

Bei der Auswertung der Daten des Low-Budget-Settings fiel auf, dass bei sechs von acht Testpersonen eine Auswertung über die Kopfkamera nicht möglich war. Die Kopfbewegungen waren zu schwach und

² Die Applikation MultiTimer kann kostenlos unter <https://github.com/GrischaSchmiedl/Multitimer> geladen werden.

/ oder zu kurz. Es war entweder keine sichere Messung möglich oder es gingen bis zu 44 % der Interaktionseinheiten verloren. Die Verwendung der zusätzlichen frontal auf die Testperson gerichtete Kamera war daher unabdingbar, wobei die Synchronisation der beiden Videoaufnahmen bei der Auswertung zu zusätzlichem Arbeitsaufwand führt.

High-Budget

Sofern die Kalibrierung des mobilen Eyetrackings erfolgreich war, erreichte die Auswertung der Daten des mobilen Eyetrackings annähernd die gleichen Ergebnisse der Auswertung von Kamera 2 (vgl. Abbildung 1: Hardware Setup). Die Abweichungen lagen bei der Anzahl der Interaktionseinheiten bei 1, bei der durchschnittlichen Dauer einer Interaktionseinheit bei 1/10 Sekunde und bei 1,7 Sekunden bei der Gesamtdauer der Interaktionseinheiten.

In unserer Testreihe war es uns nur bei einer Testperson gelungen, das mobile Eyetracking so gut zu kalibrieren, dass die Position des Blickes auf das Handy hinreichend genau bestimmt werden konnte (siehe Abbildung 4).



Abbildung 4: Blickpunktbestimmung beim mobilen Eyetracking

7 Fazit

Beim *Low-Budget-Setting* sind unbedingt zwei Kameras nötig. Die Kopfkamera könnte jedoch durch eine fixe Kamera über der Schulter ersetzt werden. So ist es möglich, über Kamera 2 die Interaktionseinheiten zu bestimmen und über Kamera 1 die Interaktion auf dem Smartphone zu verfolgen. Dadurch können Beeinträchtigungen des Users durch die Kopfkamera vermieden werden. Andererseits kann es sich als nachteilig herausstellen, dass das Geschehen des Simulators nicht mehr verfolgt werden kann und damit Reaktionen des Users auf den Simulator eventuell nicht erklärt werden können.

Beim *High-Budget-Setting* könnte auf Kamera 2 verzichtet werden. Wir raten jedoch davon ab. Ohne diese Kamera ist es nicht möglich das Gesicht der Testperson zu erfassen und somit die Gefühlsregung der Person zu erkennen. Außerdem ist bei dem von uns verwendeten mobilen Eyetracking-System kein Mikrophon integriert. Der Ton muss separat aufgezeichnet werden. Dies kann über Kamera 2 erfolgen.

Die Probleme des mobilen Eyetrackings liegen in der Kalibrierung. Sie muss für jede Testperson neu durchgeführt werden. Dies erhöhte in unseren Versuchen den Zeitaufwand für die Einrichtung des technischen Systems um durchschnittlich 273 Prozent. In unserem Test waren sechs der acht Testpersonen Brillenträger. Vier davon konnten problemlos eingerichtet werden. Bei zwei Testpersonen mit Brille war keine Kalibrierung möglich. Über die Eignung bestimmter Brillentypen für das mobile Eyetracking kann jedoch keine Aussage getroffen werden. Sofern das Eyetracking richtig auf die Testperson kalibriert ist, ist eine genauere Blickerfassung möglich. Bei unserem Test konnten wir jedoch keine zusätzlichen Daten ermitteln, bspw. über die Zeit, die ein User für das Auffinden eines Buttons benötigt.

Mit dem von uns getesteten mobilen Eyetracking lässt sich kein ausreichender Mehrwert gegenüber dem Low-Budget-Setting erzielen, der die Kosten für die teurere Lösung rechtfertigen würden. Zusätzliche Daten stehen im Vergleich zum Low-Budget-Setting zur Auswertung nicht zur Verfügung. Die *Blickverlaufsmessung* – der eigentliche Mehrwert eines Eyetracking-Systems in unserem Setting – ist durch das kleine Display des Smartphones relativ ungenau und lässt nicht immer das Erkennen des genauen Blickfokus zu. Außerdem ist die typische *Blickdauer*³ in den einzelnen Interaktionen mit dem Gerät so kurz, dass keine Erkenntnisse über Probleme bei der Gestaltung des User Interfaces gewonnen werden konnten.

³ Die typische Blickdauer ist abhängig vom eingesetzten Primärszenario. Sie lag bei unserer Autosimulation bei durchschnittlich 1,2 Sekunden.

Jedoch zeichnen sich im Bereich des mobilen Eyetrackings mit Eyetracking-Brillen⁴ Fortschritte ab. Die neuesten Entwicklungen der beiden größten Eyetracking-Hersteller Tobii und SMI versprechen bessere und genauere Messungen sowie Auswertung auch bei mobilen Geräten und in der Bewegung. Die Kosten für ein High Budget Testsetting mit der neuen Eyetracking-Brille belaufen sich allerdings auf mindestens 27.000 Euro (ohne Simulator).

8 Weiterführende Forschung

Um Daten genauer auswerten zu können, wäre die Entwicklung eines Logbuches denkbar. Hier könnten alle Daten, die der User durch seine Interaktion auslöst, verfolgt werden. Im Abgleich mit den Videoaufzeichnungen sind genauere und womöglich mehr Daten erfassbar.

Zudem sind andere Spielekonsolen (bspw. Xbox 360 mit Kinect-Kamera) und Spiele denkbar, um den Grad der Ablenkung des Primärtasks veränderbar zu gestalten. Damit könnten unterschiedliche reale Szenarien der mobilen Nutzung hinsichtlich der maximalen Zeit einer Interaktionseinheit besser simuliert werden.

9 Literaturverzeichnis

Coursaris, C.K. & Kim, D. (2007) A research agenda for mobile usability. In: *CHI '07 extended abstracts on Human factors in computing systems*. CHI '07. New York, NY, USA, ACM, S.2345–2350.

Kallio, T. & Kaikkonen, A. (2005) Usability testing of mobile applications: A comparison between laboratory and field testing. *Journal of Usability Studies*, 1, S.4–16.

Kim, H., Kim, J., Lee, Y., Chae, M. & Choi, Y. (2002) An Empirical Study of the Use Contexts and Usability Problems in Mobile Internet. In: *PROCEEDINGS OF THE 35 TH ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, LOS ALAMITOS, CA: IEEE COMPUTER*. Available from: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.98.7561>> [Zugegriffen 28 Januar 2011].

Kristoffersen, S. & Ljungberg, F. (1999) “Making place” to make IT work: empirical explorations of HCI for mobile CSCW. In: *Proceedings of the international ACM SIGGROUP conference on Supporting group work*. New York, NY, USA, ACM Press, S.276-285. Available from: <<http://portal.acm.org/citation.cfm?id=320297.320330&coll=DL&dl=GUIDE&CFID=7827411&CFTOKEN=90559788>>.

Looije, R., te Brake, G.M. & Neerinx, M.A. (2007) Usability engineering for mobile maps. In: *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*. Mobility '07. New York, NY, USA, ACM, S.532–539.

Oulasvirta, A. & Nyssönen, T. (2009) Flexible Hardware Configurations for Studying Mobile Usability. *Journal of Usability Studies*, 4 (2), S.93-105.

Oulasvirta, A., Tamminen, S., Roto, V. & Kuorelahti, J. (2005) Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. CHI '05. New York, NY, USA, ACM, S.919–928.

Perry, K.B. & Hourcade, J.P. (2008) Evaluating one handed thumb tapping on mobile touchscreen devices. In: *GI '08 Proceedings of graphics interface 2008*. GIC '08. Windsor, Ontario, Canada, ACM, S.57-64.

Schmiedl, G., Blumenstein, K. & Seidl, M. (2011) Usability Testing for Mobile Scenarios of Fragmented Attention. In: *Proceedings of the CHI Sparks Conference, HAN*. Arnhem, Nederland.

Schusteritsch, R., Wei, C.Y. & LaRosa, M. (2007) Towards the perfect infrastructure for usability testing on mobile devices. In: *CHI'07 extended abstracts on Human factors in computing systems*. S.1839–1844.

Zhang, D. & Adipat, B. (2005) Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. *International Journal of Human Computer Interaction*, 18 (3), S.293-308.

⁴ Eyetracking Glasses von SMI (<http://eyetracking-glasses.com/>) und Tobii Glasses (<http://www.tobii.com/eye-tracking-research/global/products/hardware/tobii-glasses-eye-tracker/>)