

Formative Usability-Evaluation der Smartphone-App (OFFTIME)

Jette Beißer

(Matrikel-Nummer: 539592)

Humboldt-Universität zu Berlin

Bachelorarbeit (8. Semester)

Ingenieurpsychologie

Betreuung: Prof. Hartmut Wandke; Prof. Annekatrin Hoppe

Wortanzahl:

Zusammenfassung: 156; Text: 8.297

Referenzen: 32

Korrespondenz zu diesem Manuskript an jette.beisser@googlemail.com

Inhaltsverzeichnis

Zusammenfassung	4
1. Einleitung.....	5
1.1. Evaluationsziele	9
2. Methoden.....	9
2.1. Heuristische Evaluation.....	9
2.2. Formative Evaluation	10
2.2.1. Stichprobe	10
2.2.2. Untersuchungssituation.....	10
2.2.3. Aufgaben.....	12
2.2.4. Variablen	13
2.2.5. Datenanalyse	13
2.3. Summative Evaluation	13
2.3.1. Variablen	14
2.3.2. Datenanalyse	14
3. Ergebnisse.....	15
3.1. Heuristische Evaluation.....	16
3.2. Formative Evaluation	18
3.2.1. Erster Testzyklus	18
3.2.1.1. Erwartungen.....	19
3.2.1.2. Ergebnisse.....	19
3.2.1.3. Interpretation.....	21
3.2.1.4. Designimplikationen.....	21
3.2.2. Zweiter Testzyklus	23
3.2.2.1. Erwartungen.....	23

3.2.2.2.	Ergebnisse	24
3.2.2.3.	Interpretation.....	25
3.2.2.4.	Designimplikationen.....	26
3.2.3.	Dritter Testzyklus	27
3.2.3.1.	Erwartungen.....	28
3.2.3.2.	Ergebnisse.....	28
3.2.3.3.	Interpretation.....	29
3.2.4.	Iterationsübergreifende Ergebnisse	29
3.3.	Summative Evaluation	30
4.	Diskussion	34
Literaturverzeichnis.....		38
Anhang		42
Selbstständigkeitserklärung.....		48

Zusammenfassung

(OFFTIME) – Eine Smartphone-App, deren Aufgabe es ist, den Alltag ihrer Nutzer in einer hypervernetzten Welt zu strukturieren. Am Arbeitsplatz verlockende Spiele-Apps sperren und nach Feierabend nicht mehr für die Kollegen erreichbar sein. In Phasen der Konzentration die Selbstkontrolle und in der Freizeit die Entspannung unterstützen. Erreichen kann (OFFTIME) diese Ziele allerdings nur, wenn es den Nutzern einfach gemacht wird, die helfenden Barrieren zu errichten. Deshalb wurde in der vorliegenden Studie versucht, mithilfe eines entwicklungsbegleitenden, formativen Evaluationsprozesses die Usability der App zu steigern. In drei Iterationen wurden schrittweise angepasste (OFFTIME)-Prototypen mit alltagsnahen Aufgaben untersucht. Die hierbei integrierten Veränderungen des Designs der Prototypen erfolgten unter Zuhilfenahme der Prinzipien nach Norman (1988) und Heuristiken nach Nielsen (1995). In vielen Bereichen konnten Verbesserungen der Effektivität, Effizienz und vor allem der Zufriedenheit der Nutzer mit (OFFTIME) entsprechend der ISO 9241-11 (1999) erreicht werden. Weiterer Optimierungsbedarf besteht vor allem beim strukturellen Aufbau der Einstellungsmenüs.

1. Einleitung

Ständig verfügbar über Handy und Internet sind Zeit-Reisende der Gegenwart mit allem Möglichen in Kontakt – bloß nicht mit sich selber. Zu gehetzt, zu leistungsgetrieben, um bei sich anzukommen. Fremdbestimmt.
Angela Elis (2011), Spiegel-Online

Handys und Smartphones sind zu einem festen, wichtigen Bestandteil des modernen Alltags geworden. Im Frühjahr 2014 war in den USA und in Deutschland etwas mehr als jeder zweite Einwohner Nutzer eines Smartphones (Statista, 2014) – also eines Mobiltelefons mit erweitertem Funktionsumfang wie E-Mail, Navigation, Ein- und Ausgabe von Bild- und Tonmaterial, sowie einem komplexen Betriebssystem (Gabler Wirtschaftslexikon, 2011). Die Erfahrung und der damit einhergehende Druck, dauerhaft auf verschiedensten Kommunikationswegen erreichbar zu sein, sind dadurch für viele Menschen allgegenwärtig geworden. Das Resultat sind unter anderem Schwierigkeiten bei der Trennung von Privatem und Berufsleben, Stress oder sogar Schlafstörungen (Thomée, 2012; Yun, Kettinger & Lee, 2012). Abschalten – im wahrsten Sinne des Wortes – fällt vielen Menschen dennoch schwer. Das Gefühl, etwas zu verpassen, sich in seinen Nutzungsmöglichkeiten selbst zu beschneiden und aus dem gewohnten Umfeld der permanenten Erreichbarkeit auszuschließen, kann ebenfalls Stress und Unwohlsein hervorrufen (Västra Götalandsregionen, 2012, zitiert nach Grauers und Wall, 2012). Nicht umsonst sind deshalb auch in Schulklassen, bei Vorlesungen, bei der Arbeit und sogar beim Autofahren immer wieder Menschen zu sehen, die ihre Finger und Augen nicht von den Allzweck-Telefonen lassen können.

Genau an diesem Punkt wollten die Entwickler der App (OFFTIME) anknüpfen, als sie sich 2013 zusammenfanden, und ihr Startup-Unternehmen gründeten. Den Nutzern ihrer Smartphone-App soll die Möglichkeit geboten werden, sich einfacher zu disziplinieren. Sie sollen darüber nachdenken, welche Anwendungen, welche Funktionen, welche Kontakte sie in unterschiedlichen Lebenslagen wirklich brauchen. Ablenkungen sollen im hypervernetzten Alltag reduziert werden, um konzentriertes Arbeiten zu erleichtern, aber auch Entspannungsphasen zu unterstützen. Detaillierte Einstellungsfunktionen ermöglichen es,

differenzierte Profile zu erstellen, die allen Alltagssituationen angepasst sind. Die Wirkung, die (OFFTIME) ihren Nutzern dadurch verspricht, ist eine Reduktion von Stress, die Steigerung von Konzentration und Produktivität, eine adaptive Einbindung des Smartphones in die aktuellen, individuellen Bedürfnisse.

Die vorliegende Studie befasst sich mit einer der primären Voraussetzungen dafür, dass diese Ziele auch tatsächlich erreicht werden können. Die versprochene Wirkung von (OFFTIME) kann nur eintreten, wenn es den Nutzern möglich ist, die Funktionsweise der App zu verstehen, ihre Funktionen zu entdecken und die gewünschten Einstellungen vorzunehmen. Deshalb galt es, die Usability der App mithilfe einer Usability-Evaluation Schritt für Schritt zu steigern.

Usability (Gebrauchstauglichkeit) beschreibt dabei, inwieweit ein Produkt, im vorliegenden Fall eine Software, es seinen Nutzern ermöglicht, gewünschte Ziele effektiv, effizient und zufriedenstellend zu verwirklichen. Unter Effektivität wird in diesem Kontext verstanden, ob Nutzer bestimmte Ziele vollständig und genau erreichen können. Dahingegen bezieht sich die Effizienz auf den Aufwand, den die Nutzer betreiben müssen, um ebendiese Ziele mit dem Produkt zu erreichen. Die Zufriedenheit letztlich beschreibt, ob die Nutzer eine positive Einstellung gegenüber dem Produkt aufweisen und es ohne ein Gefühl von Beeinträchtigung verwenden können (ISO 9241-11, 1999). Die Untersuchung der Usability beschäftigt sich demnach mit einem vorgelagerten Problem, welches betrachtet werden muss, noch bevor die arbeits- und leistungs-, aber auch entspannungsbezogenen Gewinne der App zum Tragen kommen können.

Um die Usability der App zu steigern, wurde in dieser Studie im Sinne der partizipativen Softwareentwicklung (Rauterberg, 1991) die Methode der formativen Evaluation gewählt. Hierbei handelt es sich um ein iteratives, entwicklungsbegleitendes Vorgehen (Lewis, 2014; Gould, 1988). Während des Entwicklungsprozesses wird die Software dabei immer wieder mit einer kleinen Anzahl von Personen aus der Zielgruppe getestet. Es soll herausgefunden werden, welche Verbesserungen im Produkt zu integrieren sind und ob deren Umsetzung die gewünschten Resultate nach sich zieht. Dadurch wird versucht, noch vor Beendigung der Entwicklungsarbeit die Usability des Produktes zu optimieren (Hegner, 2003). Dieser Prozess des Durchlaufens von mehreren Iterationen mit jeweils etwa fünf Versuchspersonen wird auch von Nielsen (2000, 2011) als derjenige beschrieben, bei dem das Verhältnis von Kosten, Aufwand und Nutzen am günstigsten ist.

Jede Iteration umfasst dabei eine Usability-Testung, die die Entwicklung eines neuen Prototypen nach sich zieht. Dieser wiederum geht als Ausgangsprodukt in die folgende Iteration ein (Shackel, 1991). In diesem Zusammenhang wird stets die große Bedeutung von qualitativen Daten betont. Denn es sei an diesem Punkt der Entwicklung wichtiger zu wissen „why the interface is wrong than how much it is wrong“ (Nielsen, 1992, S.18).

Zudem sind formative Evaluationen ein wichtiger Aspekt beim Prozess des Human-Centered Design. Das Ziel der Entwickler ist hierbei ein Produkt zu schaffen, welches es dem Nutzer ermöglicht, auf seinem bestmöglichen Leistungsniveau zu agieren. Das Produkt soll dabei die Wünsche und Bedürfnisse des Nutzers erfüllen und keine Anpassung des Nutzers an das System – zum Beispiel durch Training – erfordern (Greenhouse, n.d.).

Um das Ziel der gesteigerten Usability zu erreichen, mussten für die Testungen Aufgaben entwickelt werden, die drei Ansprüchen gerecht werden. Erstens sollten sie repräsentativ für die alltägliche Nutzung der (OFFTIME)-App sein. Es wurde also nach problemorientierten Aufgaben im Use-Szenario (Hegner, 2003) gesucht. Dafür wurde in Zusammenarbeit mit dem Entwicklerteam erläutert, was die wichtigsten Funktionen der App und bereits bekannte Bedürfnisse der Nutzer sind. Zweitens sollten die Aufgaben verschiedene Schwierigkeitsgrade innehaben und die Funktionen in unterschiedlicher Korngröße abbilden. Dies würde einerseits die Untersuchung einzelner Bedienelemente, aber auch größerer Funktionseinheiten wie Menüstrukturen ermöglichen. Außerdem sollte durch vereinzelte, sehr einfache Teilaufgaben einer möglicherweise aufkeimenden Frustration entgegengewirkt werden. Drittens wurden die Aufgaben so gewählt, dass die Versuchspersonen mit Elementen der App interagieren, die nach Expertenevaluation Nutzungsprobleme verursachen könnten (Nielsen, 1993). Um dieses Kriterium zu erfüllen, wurde zu Beginn der Studie eine heuristische Evaluation durchgeführt. Hierbei sucht ein Experte nach Schwachstellen des Designs, die es weiter zu untersuchen oder zu beseitigen gilt (Nielsen, 1993). Diese Schwachstellen der Usability wurden mithilfe von Richtlinien und Heuristiken nach Norman (1988) und Nielsen (1995) identifiziert.

Norman (1988) beschreibt vier wichtige Prinzipien, die es dabei zu beachten gilt:

- Sichtbarkeit – Bedienelemente und Funktionen müssen klar wahrzunehmen sein
- Mapping – klare Beziehungen zwischen Bedienelementen, deren Funktionen und Auswirkungen auf das System sollen erkennbar sein
- Feedback – Rückmeldung des Systems über Ausführung und Resultat einer Aktion
- Gutes konzeptuelles Modell – ermöglicht Vorhersage von Auswirkungen einer Aktion/Handlung auf das System, sowie bessere Planung zielführender Aktionen

Die zehn Heuristiken nach Nielsen (1995) erweitern und ergänzen diese Prinzipien:

- Sichtbarkeit des Status des Systems
- Übereinstimmung von System und realer Welt – etwa Nutzung von Zeichen-Konventionen, Einhaltung der Leserichtung
- Kontrolle und Freiheit – Wiederholen/rückgängig machen von Aktionen ermöglichen
- Konsistenz und Standards – Einhaltung von Konventionen, Verwendung konsistenter Begrifflichkeiten/Anzeigen
- Fehler verhindern
- Erkennen statt Erinnern – Gedächtnisbelastung des Nutzers minimieren
- Flexibilität und Effizienz – Anpassungsmöglichkeiten für neue und Experten-Nutzer/Shortcuts
- Ästhetisches und minimalistisches Design – auf irrelevante Informationen verzichten
- Hilfe bei Erkennung und Beseitigung von Fehlern – Art der Fehler erkennbar machen, Lösungsmöglichkeiten aufzeigen
- Hilfe und Dokumentation

Die abgeleiteten Aufgaben sind im Wortlaut in Anhang 3) nachzulesen.

Zusammenfassend ist festzuhalten, dass in der vorliegenden Studie zwei unterschiedliche qualitative Evaluationsmethoden verwendet wurden. Sie können nach Hamborg & Gediga (2006, zitiert nach Kaspar, 2010) in die Kategorien der prädiktiven und deskriptiven Methoden eingeordnet werden. Bei prädiktiven Evaluationsmethoden, auch Inspektionsmethoden genannt, handelt es sich zumeist um solche, bei denen ein Usability-Experte Schwachpunkte des Produktes aufzudecken versucht. Deskriptive Methoden hingegen versuchen, die Usability eines Produktes mit Hilfe repräsentativer Benutzer zu steigern. Diese Methoden bedienen sich also einer direkten Interaktion zwischen Personen der Zielgruppe und dem untersuchten Produkt (Hegner, 2003). Der häufig als ideal angesehene Weg zur Entwicklung eines (Software-)Produktes mit hoher Usability besteht darin, beide Ansätze miteinander zu verknüpfen (Nielsen, 1992; Hartson, Andre & Williges, 2001). Aus diesem Grund wurde auch in der vorliegenden Studie die Integration beider Evaluationsformen gewählt. Die zusätzliche Betrachtung der Veränderung von Effizienz und Effektivität über die Iterationen hinweg erlaubte es, im Rahmen einer summativen Evaluation,

vorsichtige Aussagen über die quantitative Veränderung der Usability zu treffen (Nielsen, 1992; Hegner, 2003).

1.1. Evaluationsziele

Die Evaluationsziele der Studie müssen aufgrund der beschriebenen Evaluationsansätze in drei Teilen betrachtet werden. In den ersten beiden Fällen handelte es sich um analysierende Evaluationen.

Das Ziel der heuristischen Evaluation war es, geeignete Aufgaben für die eigentliche Testphase zu entwickeln (Nielsen, 1993).

Ziel der formativen Evaluation während des Iterationsprozesses war es, repräsentativen Nutzern die Möglichkeit zu bieten, direkt mit dem Produkt zu interagieren. Auf diesem Weg sollten Erkenntnisse über Schwachstellen des Produktes gesammelt, und daraus resultierende Implikationen für das Design direkt abgeleitet werden (Hegner, 2003).

Durch die zusätzliche Betrachtung von Effektivität und Effizienz im dritten Teil der Evaluation, konnten vorsichtige Aussagen über die objektiven Kennwerte von Usability nach der ISO 9241-11 getroffen werden.

2. Methoden

2.1. Heuristische Evaluation

Für die heuristische Evaluation fand zunächst ein Gespräch zwischen Entwicklern und Versuchsleiterin statt, bei dem die Anwendungsgebiete und bereits durch das (OFFTIME)-Team identifizierte Nutzer-Ansprüche besprochen wurden. Hierdurch sollten die Funktionsfelder konkretisiert werden, auf die bei der Evaluation besonderen Wert gelegt werden sollte und die es später bei der Entwicklung der Testaufgaben zu berücksichtigen galt.

Anschließend wurde unter Zuhilfenahme der Prinzipien von Norman (1988) und der Heuristiken nach Nielsen (1995) das Design geprüft und nach Stellen gesucht, an denen die Kriterien nicht oder nicht hinreichend erfüllt wurden. Auf Grundlage dieser Erkenntnisse wurden die Aufgaben entwickelt, die später in der formativen Evaluation verwendet wurden.

2.2. Formative Evaluation

2.2.1. Stichprobe

Die Versuchspersonen konnten mithilfe von Aufrufen über universitätsinterne E-Mail-Verteiler, in dem psychologischen Experimental-Server PESA, sowie in sozialen Netzwerken gewonnen werden.

Der Aufruf ist im Wortlaut in Anhang 1) nachzulesen.

Je Testzyklus gingen 5 Versuchspersonen in die Untersuchung ein. Eine Testung von 5 Nutzern je Iteration wurde gewählt, da dies der von Nielsen (2000) vorgeschlagene Umfang ist, mit dem 85% aller Usability-Probleme aufgedeckt werden können.

Das Durchschnittsalter lag bei 25,7 Jahren (1. Testzyklus: $M = 27$ ($SD = 12,4$); 2. Testzyklus: $M = 24,2$ ($SD = 6,3$); 3. Testzyklus: $M = 25,8$ ($SD = 1,8$)). Von den insgesamt 15 ausgewerteten Personen waren 9 männlich und 6 weiblich.

Die Stichprobe bestand aus Personen, die nach eigener Einschätzung intensive Nutzer eines Smartphones sind. Sie gaben im Durchschnitt an, 10,7 Apps in regelmäßigem Gebrauch zu haben (1. Testzyklus: $M = 10,2$ ($SD = 4,7$); 2. Testzyklus: $M = 17,2$ ($SD = 10,5$); 3. Testzyklus: $M = 4,8$ ($SD = 2,9$)). Da es sich bei der zu testenden App um eine Android-spezifische Software handelt, kamen entsprechend nur Nutzer von Android-Smartphones als Probanden in Frage. Außerdem hatte keine der Versuchspersonen bereits Vorerfahrungen mit der zu testenden App.

Eine Versuchsperson musste im dritten Testzyklus von der Auswertung ausgeschlossen werden. Nach Schwierigkeiten beim Lösen der ersten Aufgabe entwickelte sie ein stark reaktantes und zeitweise aggressives Verhalten, welches die instruktionskonforme Durchführung der Aufgaben unmöglich machte. Aus diesem Grund wurde eine zusätzliche sechste Person getestet, deren Performanz dann in die Auswertung einging.

2.2.2. Untersuchungssituation

Die Untersuchungen fanden als Einzeltestungen teilweise in einem Labor des Lehrstuhls für Ingenieurpsychologie der Humboldt-Universität zu Berlin statt. Da es sich bei den Versuchspersonen häufig um berufstätige Personen handelte, wurden einige Testungen auch

aus logistischen Gründen in Privatwohnungen durchgeführt. In allen Fällen wurde darauf geachtet, dass die Testungen in einer ruhigen Umgebung stattfanden, in der Ablenkung und Geräuschkulisse auf ein Minimum reduziert waren.

Die Versuchspersonen nahmen jeweils an einem Tisch Platz, auf dem sich ein Sony Xperia C1905-Smartphone befand. Auf dem Smartphone waren zuvor die zu untersuchende App (OFFTIME), sowie ein Kalender und einige Dummy-Apps und -Kontakte installiert worden. Diese Einstellungen waren notwendig, um die Testaufgaben zu lösen. Außerdem befand sich auf dem Tisch eine Kamera, mit der die Testungen aufgezeichnet wurden. Die Videos dienten der Vereinfachung der Datenauswertung, sowohl für die qualitativen, als auch die quantitativen Ergebnisse, die unter anderem in Form von Klick-Zahlen erhoben wurden.

Die Versuchsleiterin nahm schräg gegenüber der Versuchsperson Platz, um Notizen schreiben und die Kamera überprüfen zu können, sowie einen freien Blick auf den Smartphone-Bildschirm zu haben. Diese Anordnung kann in Abb. 1 nachvollzogen werden.

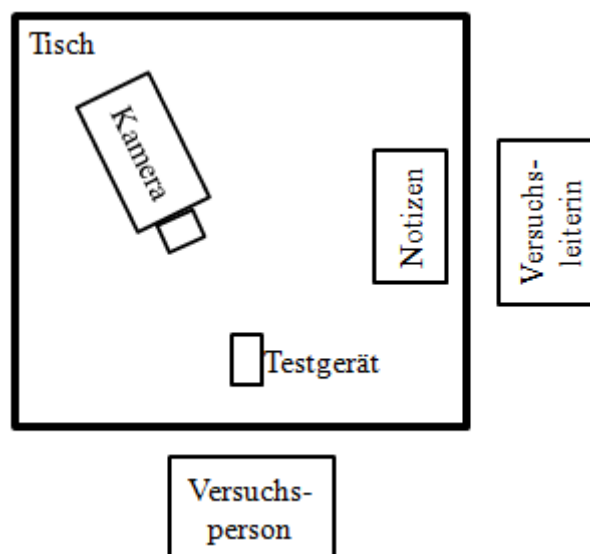


Abb. 1: Überblick über die Versuchsanordnung

Nach der Begrüßung der Versuchsperson wurde sie über den Sinn der App, sowie das Ziel der Studie, also die Verbesserung der Usability von (OFFTIME), aufgeklärt. Daraufhin wurde das Prinzip des lauten Denkens erklärt, das die Versuchspersonen im Folgenden während der gesamten Testung anwenden sollten. Bei dem sogenannten „Think-Aloud Protocol“ werden die Versuchspersonen gebeten, jeden Gedanken, den sie haben, laut zu äußern und zu beschreiben (Krug, 2010). Außerdem wurden die fünf Schritte erläutert, in

denen jede Aufgabe abgearbeitet werden sollte. Die Versuchspersonen wurden darum gebeten, jede neue Anzeige genau zu beschreiben und Bedienelemente zu identifizieren, noch bevor Vermutungen über benötigte Elemente und deren Reaktionen auf ihre Betätigung aufgestellt werden. Erst danach sollten die geplanten Handlungen zur Aufgabenerfüllung ausgeführt, und deren Ergebnisse im letzten Schritt evaluiert werden. Diese Abfolge wurden in Anlehnung an die Technik des cognitive walkthrough (Rieman, Franzke & Redmiles, 1995), sowie die CE+ Theorie nach Polson und Lewis (1990) entwickelt. Hier ist das Ziel, walk-up-and-use Systeme so zu gestalten, dass sie spontan, ohne formelles Training und Anleitung genutzt werden können. Der Nutzer suche laut Theorie das System zuerst nach aktuell verfügbaren Aktionsoptionen, etwa Bedienelemente, Menüleisten etc., ab. Im Anschluss werde zur Ausführung diejenige Aktion ausgewählt, von der der Nutzer glaube, sie könne ihn am ehestem seinem Ziel näherbringen. Erst dann werde diese Aktion ausgeführt und am Ende evaluiert, ob das erwünschte Ergebnis eingetroffen ist und man sich einen Schritt hin zur Lösung der Aufgabe begeben hat.

Die Prozedur der fünfschrittigen Aufgabenbearbeitung und das dabei verwendete laute Denken wurden bei einer Übungsaufgabe in der Kalender-App, also noch ohne Einsatz von (OFFTIME), trainiert. Bevor es dann an die Bearbeitung der tatsächlichen Testaufgaben ging, hatten alle Versuchspersonen die Möglichkeit, Fragen zum Testablauf mit der Versuchsleiterin zu klären.

Das genaue Skript für die Einweisung in die Testprozedur ist im Wortlaut in Anhang 2) nachzulesen.

2.2.3. Aufgaben

Nach dem Trainieren der Technik des lauten Denkens und der fünfschrittigen Aufgabenbearbeitung mit der Übungsaufgabe, wurden jeder Versuchsperson sechs Aufgaben gestellt, die es mithilfe von (OFFTIME) zu lösen galt. Diese Aufgaben veränderten sich während der Iterationen nicht. So konnte untersucht werden, ob die schrittweisen Veränderungen des Designs die Nutzung der App wie erhofft einfacher, fehlerfreier und angenehmer gestalteten.

Die Aufgaben unterschieden sich hinsichtlich ihres Schwierigkeitsgrades und dem nötigen Aufwand, um sie vollständig zu lösen. Zu Beginn waren sie darauf ausgerichtet, einzelne Funktionen der App zu verwenden und kennenzulernen. Dadurch konnten Bedienelemente

wie die Uhr zur Einstellung der (OFFTIME)-Dauer oder der „Zauberstab“ (später „Zahnrad“) zum Eintritt in das sequentielle Menü speziell und intensiv untersucht werden. Den Abschluss bildete eine vollständige Simulation der Einrichtung eines neuen Profils mit diversen Einstellungen, die über beide Menüstrukturen der App vorgenommen werden mussten.

Die Aufgaben sind im Wortlaut in Anhang 3) nachzulesen.

2.2.4. Variablen

Im Sinne einer unabhängigen Variablen wurden in der Untersuchung die verschiedenen Versionen (Prototypen) der (OFFTIME)-App genutzt. Nach jedem Testzyklus bestehend aus fünf Testungen wurden Teile der App angepasst, um den beobachteten Problemen der Usability zu begegnen. Die veränderte, mithilfe der Ergebnisse angepasste App-Version wurde dann als Prototyp im folgenden Testzyklus verwendet.

Qualitative Daten als Maß für die abhängige Variable, also die Zufriedenheit der Nutzer mit (OFFTIME), wurden mithilfe der fünfschrittigen Test-Prozedur und des lauten Denkens erhoben.

2.2.5. Datenanalyse

Nach jedem Testzyklus wurde eine qualitative Auswertung der Testungen durchgeführt. Es wurde nach (Teil-)Aufgaben gesucht, bei denen mehrere Versuchspersonen Schwierigkeiten bei der Lösung hatten oder über irreführende Beschreibungen und Symbole klagten. Die Erkenntnisse aus den Daten des lauten Denkens der Versuchspersonen dienten der Ableitung von Designimplikationen für die (OFFTIME)-App. Diese zu treffenden Veränderungen wurden von der Versuchsleiterin entwickelt, und in Zusammenarbeit mit dem (OFFTIME)-Entwicklerteam in eine zu realisierende Form gebracht und umgesetzt. Die so entstandene aktualisierte Version der App ging in den folgenden Testzyklus ein.

2.3. Summative Evaluation

Nach Beendigung aller Testzyklen wurde im dritten Schritt eine quantitative Datenanalyse durchgeführt.

Ziel dieser summativen Evaluation war es, die Effizienz und Effektivität der App im Sinne der ISO 9241-11 über die drei Iterationen hinweg zu überprüfen.

2.3.1. Variablen

Als unabhängige Variable dienten die drei Testzyklen, über deren Verlauf hinweg die Usability untersucht werden sollte.

Die abhängigen Variablen Effizienz und Effektivität wurden durch zwei verschiedene Arten quantitativer Daten operationalisiert. Die Effizienz wurde durch das Verhältnis aus durchschnittlich und minimal benötigter Klickanzahl repräsentiert. Die Lösungshäufigkeit einer jeden Aufgabe diente dazu, die Usability im Sinne der Effektivität abzubilden.

2.3.2. Datenanalyse

Die Effizienz beschreibt, wie nah die Klickzahlen der Lösungen der Versuchspersonen an den minimalen Klickzahlen lagen. Es wurde also eine Messung der Fehlerhäufigkeit verwendet. Dazu wurden die benötigten Klickzahlen der Versuchspersonen je Testzyklus gemittelt, die die jeweilige Aufgabe korrekt lösen konnten. Um das Ergebnis weniger anfällig für Ausreißer zu gestalten, wurden Werte, die mehr als 2 Standardabweichungen vom Mittelwert entfernt lagen, durch den Mittelwert ersetzt. Mit dem modifizierten Datensatz wurde dann ein neuer Mittelwert berechnet, der für den Vergleich zwischen den Iterationen genutzt wurde. Die Effizienz berechnete sich je Aufgabe aus durchschnittlicher Klickzahl minus optimaler Klickzahl. Je kleiner also der Wert, desto größer die Effizienz einer Funktion.

Um die Klickzahl unabhängig von persönlichen Merkmalen der Versuchspersonen zu ermitteln, wurden folgende Konventionen festgelegt:

- Texteingabe = 1 Klick; Zahl unabhängig von Tippfehlern
- Scrollen = 1 Klick je Scroll-Richtung (nach oben, nach unten); Zahl unabhängig von Scroll-Vorlieben (klein- oder großschrittig)

Die Effektivität beschreibt, wie viele Personen die Aufgaben korrekt lösen konnten. Es wurde demnach eine Messung der Lösungshäufigkeiten durchgeführt. Eine Aufgabe galt als korrekt gelöst, wenn alle Einstellungen den Vorgaben der Aufgabe entsprachen und die Versuchsperson der Ansicht war, alle Einstellungen aufgabenkonform vorgenommen zu haben.

3. Ergebnisse

Um die Ergebnisse und abgeleiteten Designimplikationen nachzuvollziehen ist es notwendig, sich einen kurzen Überblick über die Struktur und das Design von (OFFTIME) zu verschaffen. In Abb. 2 ist die Struktur der App graphisch veranschaulicht. (OFFTIME) besteht grob zusammengefasst aus einem Startbildschirm, auf dem unter anderem ein Profil gewählt, die (OFFTIME)-Dauer eingestellt und eine entsprechende (OFFTIME) gestartet werden kann („Black-Screen“). Außerdem wird der Zugang zu einem Hauptmenü (3 Balken links oben) und einem Einstellungsmenü („Zauberstab“/„Zange“ rechts oben) ermöglicht. Die App verfügt über zwei verschiedene Menüstrukturen, über die Einstellungen an den Profilen vorgenommen werden können. Das sequentielle Menü („Zauberstab“) ermöglicht beim schrittweisen Durchlaufen mehrerer Screens grobe Einstellungen, das vertikale Menü („Zange“) bietet eine Liste mit allen möglichen Optionen zur Feineinstellung der Profile.

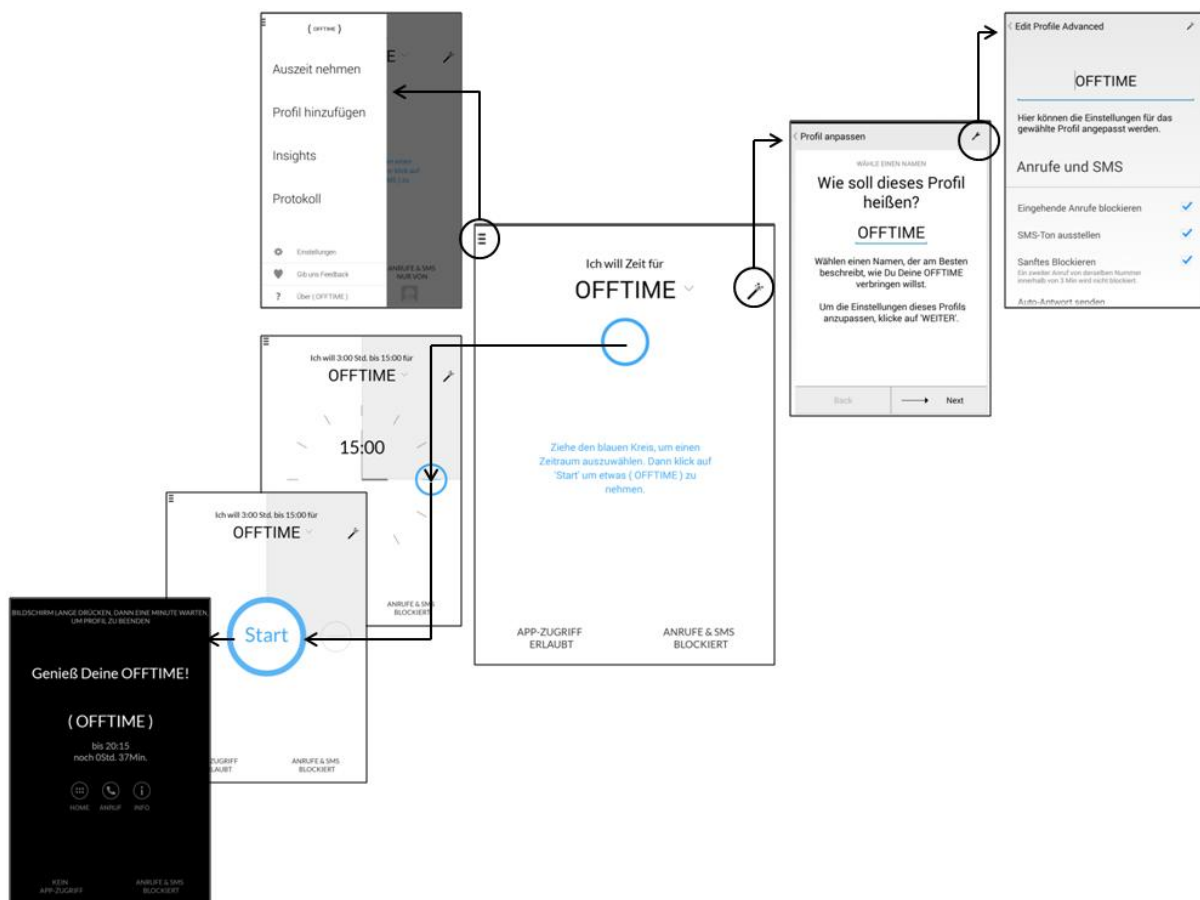


Abb. 2: Menü- und Screen-Struktur (OFFTIME)

3.1. Heuristische Evaluation

Durch die heuristische Evaluation konnten mehrere Usability-Problemfelder identifiziert werden. Vor allem die Einstellung der (OFFTIME)-Dauer über das Bedienelement „Blauer Kreis“ wurde als problematisch bewertet. Die Sichtbarkeit der Funktionsweise des „blauen Kreises“ war ebenso eingeschränkt, wie das Mapping zu seiner Funktion zur Zeiteinstellung. Außerdem war durch die sehr reduzierte Anzeige nicht mit Sicherheit gegeben, dass Nutzer das Konzept der Einstellung einer Endzeit unabhängig von der Startzeit verstehen würden. Durch das Design war es nicht ohne Hindernisse möglich, sich ein korrektes konzeptuelles Modell über die Funktion zu bilden. Auch die Erkennbarkeit der Grenzen der Funktion wurde als mangelhaft eingestuft, da nicht deutlich wurde, dass mit der Uhr lediglich eine Spanne von 12 Stunden einzustellen ist (s. Abb. 3).

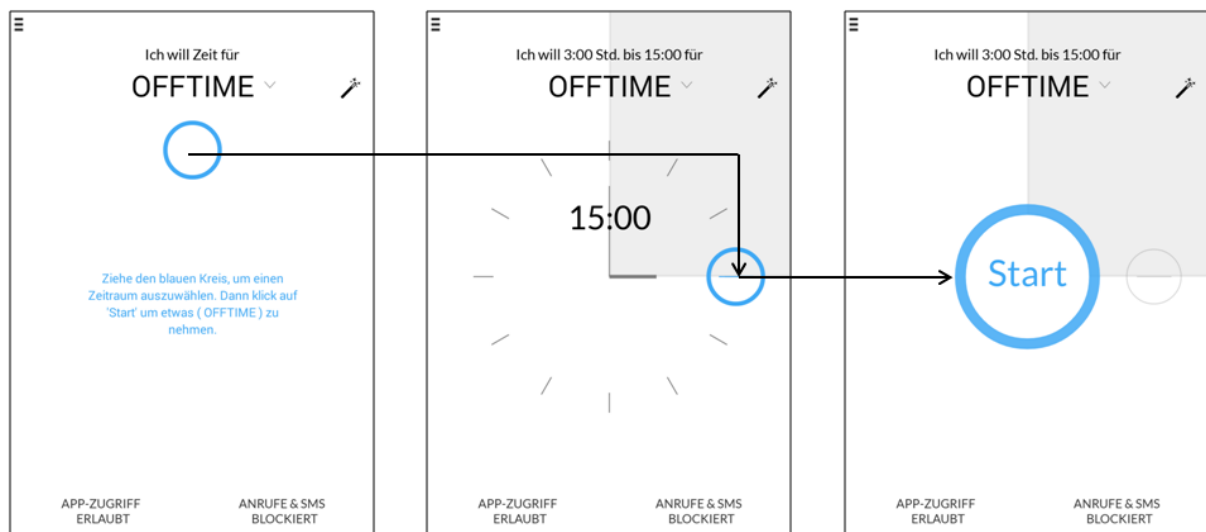


Abb. 3: Funktionseinheit zur Einstellung der (OFFTIME)-Dauer

Als sehr kritisch wurde die Verwendung neuartiger Symbole (z.B. „Zauberstab“) eingeschätzt, da hiermit das Prinzip der Nutzung von Konventionen verletzt wurde. Im Zusammenhang mit der Wahl der Symbole wurde ebenso die doppelte Struktur des sequentiellen und vertikalen Menüs als kritisch betrachtet. Es war erst durch trial & error möglich herauszufinden, ob sich beide Menüs ergänzen, doppeln oder widersprechen, und ob sie parallel oder getrennt genutzt werden müssen. Dieser Mangel an Transparenz der Struktur

kann zu Frustration und Verwirrung führen. Außerdem stellt er ein mögliches Hindernis dafür dar, den vollen Funktionsumfang der App kennenzulernen.

Generell wurde beanstandet, dass der weitreichende Verzicht auf Feedback (zum Beispiel kein „Speichern“), sowie die Verwendung von sehr minimalistischem Feedback (etwa bei der Sperrung von Kontakten nur Anzeige eines Profilbildes, keine Anzeige des Namens; s. Abb. 4) zu Schwierigkeiten in der Nutzung führen kann. Unzureichendes Feedback zieht häufig Unsicherheit hinsichtlich des Status der Einstellungen und bezüglich möglicherweise begangener Fehlern nach sich. Für den Nutzer ist es dann schwer zu erkennen, ob eine Eingabe korrekt getätigt und vom System übernommen wurde, ob sie seinen tatsächlichen Vorstellungen und Zielen entspricht.

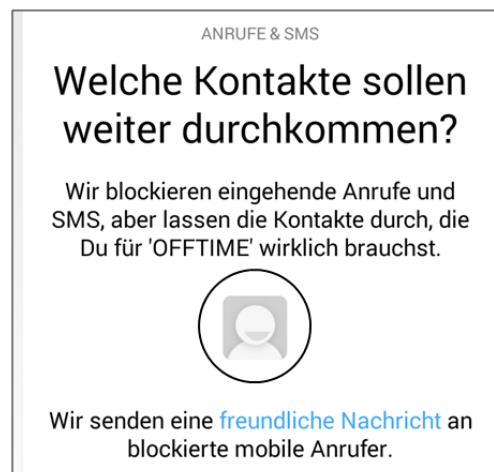


Abb. 4: Mangelndes Feedback über ausgewählte Kontakte

Es wurde eine sprachliche Inkonsistenz (Englisch, Deutsch) entdeckt, die grundsätzlich zu vermeiden ist (s. Abb. 5).

Besondere Defizite wurden zudem bei der Kalenderintegration festgestellt. Deren Aufbau war nicht klar verständlich, ergänzende Anzeigen mit Bedienungshinweisen konnten die Funktionsweise nicht hinreichend erklären. Ein Feedback in Form eines Aufleuchtens bei Klick auf einen Termin suggerierte eine Terminauswahl. Allerdings konnten Termine tatsächlich nur über die Suchfunktion ausgewählt werden. Es wurde außerdem kein klares Feedback darüber gegeben, bei welchen Terminen des Kalenders das Profil letztlich gestartet wird.

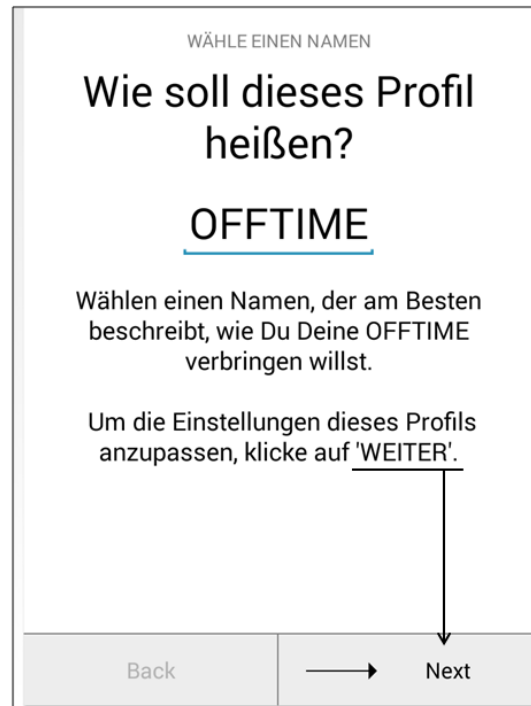


Abb. 5: Inkonsistente Nutzung von deutscher und englischer Sprache

3.2. Formative Evaluation

Mithilfe der Aufgaben, die basierend auf den Ergebnissen der heuristischen Evaluation generiert wurden, fand im Anschluss die formative Evaluation statt. In drei Iterationen wurde je ein Prototyp getestet, wurden die qualitativen Ergebnisse ausgewertet und dadurch entsprechende Designimplikationen abgeleitet. Diese konnten dann in einen neuen Prototypen integriert werden, der in den Testzyklus der folgenden Iteration einging. Die Ergebnisse der drei Testzyklen sind im Folgenden beschrieben.

Eine tabellarische Übersicht der Problementwicklung ist in Anhang 4) zu finden.

3.2.1. Erster Testzyklus

Im ersten Testzyklus diente die zu diesem Zeitpunkt aktuelle Beta-Entwicklungsversion von (OFFTIME) als Prototyp. Getestet wurde die App mit den sechs bereits beschriebenen Aufgaben aus Anhang 3).

3.2.1.1. Erwartungen

Nach der heuristischen Evaluation wurden folgende Problemfelder erwartet:

- Einstellung der (OFFTIME)-Dauer; Funktionsweise des Bedienelementes „Blauer Kreis“ schwer zu identifizieren
- Erkennen der aktuellen Start- und Endzeit
- Eintritt in sequentielles Menü; Bedeutung des „Zauberstabes“ nicht intuitiv
- Kein klares Verständnis der zwei Menüstrukturen sequentiell/vertikal als grobe, beziehungsweise Feineinstellungen
- Nutzung der Kalenderintegration; Funktionsweise des Filters nicht klar zu erkennen

3.2.1.2. Ergebnisse

Im ersten Testzyklus konnten mithilfe der qualitativen Daten drei Usability-Probleme ausfindig gemacht werden, die bei allen Versuchspersonen in sehr ähnlicher Form auftraten.

Das erste entscheidende Problem betraf die Zeiteinstellung auf dem Startbildschirm der App. Mithilfe des Bedienelementes „Blauer Kreis“, der im Uhrzeigersinn zu drehen war, sollte die gewünschte (OFFTIME)-Dauer eingestellt werden. Hierbei traten alternierend, aber auch gleichzeitig, zwei Probleme auf:

- Funktion des Bedienelementes „Blauer Kreis“ wurde nicht erkannt
- Funktionsweise des Bedienelementes „Blauer Kreis“ wurde nicht erkannt

Das zweite über alle Versuchspersonen hinweg auftretende Problem bestand im Verständnis der Funktion des „Zauberstabes“. Dieses Bedienelement ermöglichte den Zugang zu dem sequentiellen Einstellungsmenü (s. Abb. 6). Die Versuchspersonen hatten entweder keine oder nur eine sehr grobe Vorstellung davon, welche Funktion sich hinter dem Symbol verbirgt. Eine Versuchsperson dachte, es handle sich um einen interaktiven Mauszeiger, eine andere nahm an, das Symbol habe gar keine Funktion und sei lediglich ein Design-Element von (OFFTIME).

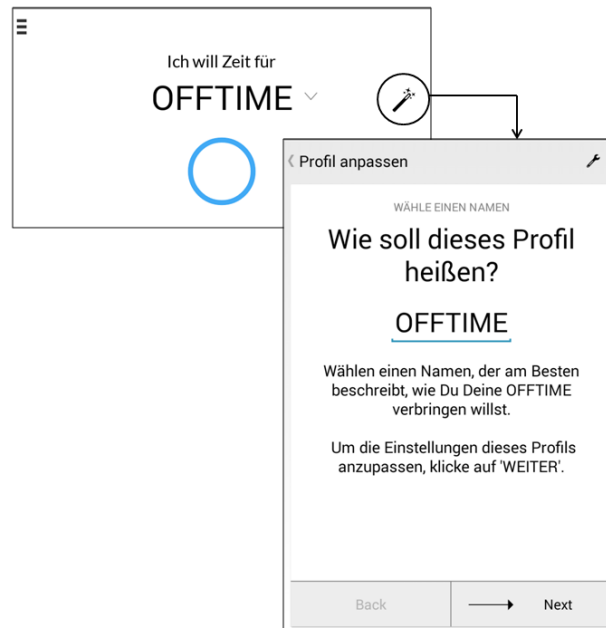


Abb. 6: "Zauberstab" zum Eintritt in das sequentielle Menü

Das dritte wichtige Usability-Problem bezog sich auf das Hinzufügen eines neuen Profils. In der App konnte nicht nur zwischen bereits existierenden Profilen gewählt, sondern auch ein neues Profil erstellt und den eigenen Bedürfnissen angepasst werden. Diese beiden Funktionen wurden in der App räumlich getrennt angeboten (s. Abb. 7). Dies entsprach nicht den Vermutungen der Versuchspersonen. Es führte dazu, dass Personen, die die Funktion der Profilwahl verstanden hatten, hier auch die Option „Profil hinzufügen“ suchten und nicht fündig wurden.

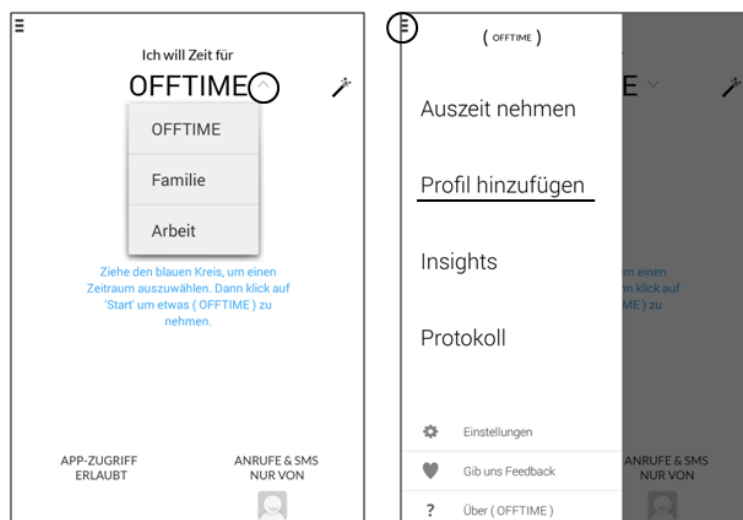


Abb. 7: Unterschiedliche Funktionen zur Wahl und zum Hinzufügen eines Profils

3.2.1.3. Interpretation

Im Fall des „blauen Kreises“ handelte es sich vorrangig um ein Mapping- bzw. Sichtbarkeits-Problem. Da das Bedienelement nur aus einem blauen Kreis bestand, wurde den Probanden nicht klar, dass es sich um ein Element zur Zeiteinstellung handelt. Eine entsprechende Beschriftung wurde nicht wahrgenommen, nicht beachtet oder nicht verstanden. Das Bedienelement als solches war sichtbar, seine Funktionalität allerdings nicht eindeutig zu identifizieren. Auch Versuchspersonen, die die grundsätzliche Funktion des Bedienelementes verstanden, konnten nicht erkennen, auf welche Art und Weise der Kreis verwendet werden sollte. Alles in allem war es den Versuchspersonen also nicht möglich, ein konzeptuelles Modell über die Funktionsweise der Zeiteinstellung mithilfe des „blauen Kreises“ zu entwickeln.

Beim „Zauberstab“ zeigte sich, dass das Symbol nicht selbsterklärend und seine Funktion nicht transparent war. Es entstand durch den Verzicht auf konventionelle Symbole auch hier ein Problem beim Mapping zwischen Symbol/Bedienelement und Funktion.

Das Problem der Wahl oder des Hinzufügens eines Profils lag daran, dass die Versuchspersonen einen semantischen Zusammenhang der Funktionen Profilwahl und Profil hinzufügen herstellten. Diese inhaltliche Beziehung wurde allerdings nicht auf das Design gemappt, also nicht durch einen räumlichen Zusammenhang repräsentiert. Dies zog Schwierigkeiten bei der Nutzung der Funktionen nach sich.

3.2.1.4. Designimplikationen

Um den drei identifizierten Usability-Problemen entgegenzuwirken, wurden im Anschluss an den ersten Testzyklus unter Beachtung der Prinzipien von Norman (1988) und Nielsen (1995) Designimplikationen abgeleitet. Diese wurden in den Prototypen der App integriert, der als Testobjekt in die folgende Iteration eingehen sollte.

Der „blaue Kreis“ war Teil des Funktionskomplexes für die Einstellung der Dauer einer (OFFTIME) und dem Starten eines Profils. Das Ziel bestand darin, diese Funktionen besser sichtbar zu machen. Dazu sollten drei Interventionen beitragen:

- Mapping zwischen Zeiteinstellung und „blauem Kreis“ herstellen, Notwendigkeit der Textbeschreibung im Sinne des minimalistischen Designs aufheben

- Status besser sichtbar machen, räumliche Trennung zwischen Zeitanzeige und Bedienelementen der Funktionsgruppe aufheben
- Grenzen der Zeiteinstellung besser sichtbar machen

Um dies zu erreichen, wurde in Zusammenarbeit mit dem (OFFTIME)-Entwicklerteam die Uhr zur Einstellung der (OFFTIME)-Dauer umstrukturiert (s. Abb. 8). Auf dem Startbildschirm war nun neben dem Bedienelement „Blauer Kreis“ permanent eine stilisierte Analoguhr eingeblendet. Der Aspekt des spielerischen Entdeckens der Funktion, der den Entwicklern besonders wichtig war, wurde trotzdem beibehalten. Nach wie vor veränderte sich die Anzeige nach dem Drücken des „blauen Kreises“. Allerdings rückte die gewählte Dauer nun in das Zentrum des Screens. Sie blieb dort auch nach Ende des Einstellungsvorgangs, anstatt, wie zuvor, lediglich am oberen Rand angezeigt zu werden.

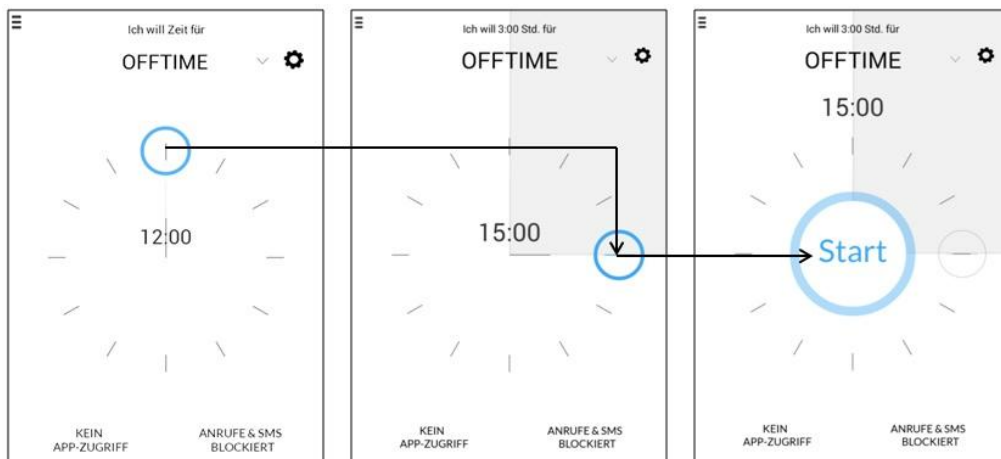


Abb. 8: Veränderte Funktionseinheit zur Einstellung der (OFFTIME)-Dauer nach Iteration 1

Der „Zauberstab“ wurde im Sinne der Einhaltung von Konventionen durch ein Zahnrad-Symbol ersetzt (s. Abb. 9). Die Verwendung eines Symbols aus der Android-Standardbibliothek ermöglichte es auf das Prinzip „Erkennen, statt Erinnern“ zurückzugreifen. Demnach ist das Erkennen (Recognition) eines bekannten Konzeptes („Zahnrad“) mit weniger Belastung der kognitiven Kapazitäten verbunden, als das Lernen und Abrufen (Recall) eines neuen Mappings („Zauberstab“) (Nielsen, 1995). An die Stelle des „Zangen“-Symbols für den Zugang zum vertikalen Menü trat das Symbol „Zahnräder“ (s. Abb. 9). Auch dieses Symbol bediente sich dem bekannten Konzept, allerdings in abgewandelter Form. Die Verbindung mehrerer Zahnräder sollte als Mapping der Hierarchie

und Hinweis darauf dienen, dass an dieser Stelle Feineinstellungen vorgenommen werden können.



Abb. 9: Neue Symbolik für Einstellungsmenüs: "Zahnrad" und "Zahnräder"

Für die Funktion des Hinzufügens eines Profils erfolgte eine Umstrukturierung der Bedienwege, um eine Konsistenz mit dem mentalen konzeptuellen Modell der Nutzer herzustellen. Die semantisch zusammenhängenden Funktionen „Profil wählen“ und „Profil hinzufügen“ sollten auch räumlich verknüpft werden. Dazu wurde im Drop-Down-Menü zur Profilwahl die Option „Profil hinzufügen“ ergänzt.

3.2.2. Zweiter Testzyklus

Im zweiten Testzyklus wurde ein Prototyp verwendet, in den die beschriebenen Designimplikationen aus Iteration eins integriert worden waren. Die Testaufgaben blieben konstant.

3.2.2.1. Erwartungen

Es wurde erwartet, dass sich nach der Integration der ersten Designimplikationen die Probleme verschieben würden. Die Veränderungen an der Zeiteinstellungsfunktion wurden von den Entwicklern zunächst vorsichtig gewählt. Deshalb wurde erwartet, dass sich die Usability dieser Funktion zwar leicht verbessert, aber weiterhin Probleme auftreten würden. Vor allem beim Erkennen der Start- und Endzeit wurden weiterhin Schwierigkeiten vermutet, ein Überblick über den Status der Zeiteinstellung würde also schwer zu erlangen sein.

Das Verständnis der Nutzer für die Zugänge zu den Menüs und deren hierarchische Struktur sollte durch die neue Symbolik hergestellt worden sein.

Zudem sollte das Hinzufügen neuer Profile nun problemlos funktionieren.

3.2.2.2. Ergebnisse

Auch im zweiten Testzyklus waren drei große Usability-Probleme klar abzugrenzen.

Die Einstellung der (OFFTIME)-Dauer auf dem Start-Screen stellte sich als nach wie vor problematisch heraus. Zwar wurde die Funktion des Bedienelementes „Blauer Kreis“ besser erkannt, die Art und Weise der Bedienung, also seine Funktionsweise, wurde aber weiterhin als nicht intuitiv eingestuft. Erwartet wurde etwa ein Klicken auf den Kreis oder eine direkte Zifferneingabe. Außerdem zeigte sich, dass die Anzeige der eingestellten Zeit noch nicht hinreichend transparent gestaltet wurde. Einige Versuchspersonen probierten, die Zeitdauer auszurechnen und zeigten sich unzufrieden darüber, dass Start-und Endzeit nicht klar zu erkennen waren.

Die zweite Problemgruppe bezog sich auf diverse Inkonsistenzen, die die Nutzung der App erschwerten. So stimmten etwa Texthinweise und Symbole nicht überein (s. Abb. 10).

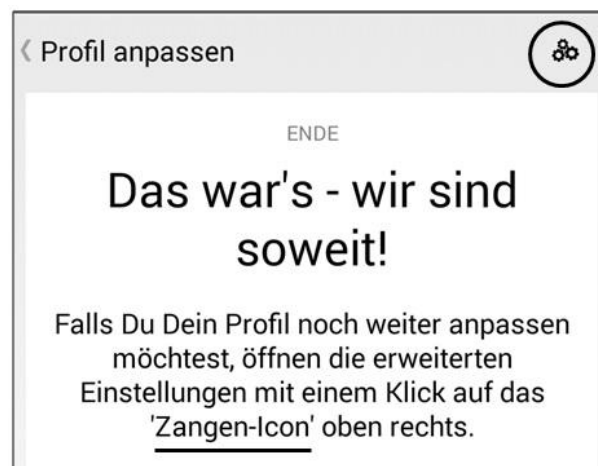


Abb. 10: Inkonsistenz zwischen Anzeigetext und Bedienelement

Auch innerhalb von eigentlich gleich strukturierten Einstellungsfunktionen (Kontakte/Apps blockieren, Ausnahmen definieren) traten Inkonsistenzen auf. Verschieden aufgebaute Suchfunktionen und Anzeigereihenfolgen sorgten für Verwirrung. So wurden zwar bei der Auswahl von Apps, die nicht blockiert werden sollten, alle installierten Apps angezeigt

und die Wahl der Ausnahmen war sehr einfach. Bei der äquivalenten Funktion für Kontakte allerdings wurde nach häufigsten Kontakten gefiltert und innerhalb dieses Filters die Suchfunktion eingesetzt. Es wurde also vorausgesetzt, dass der Nutzer weiß, in welchem Filter ein bestimmter Kontakt zu finden ist (s. Abb. 11)

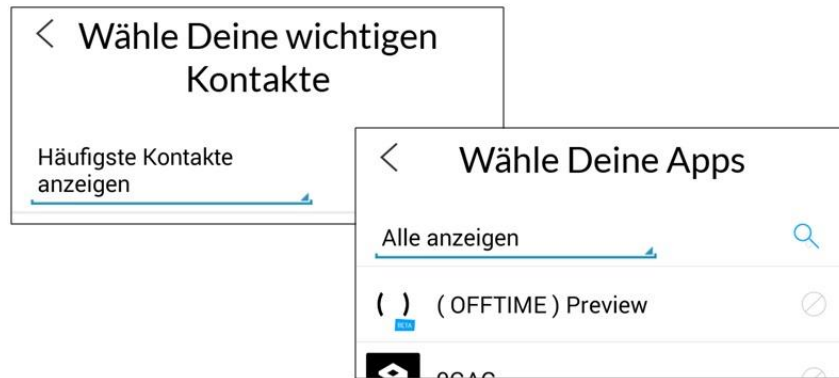


Abb. 11: Inkonsistenzen in der Struktur eigentlich äquivalenter Funktionen

Als drittes traten die Defizite der App im Geben von Feedback in den Vordergrund. Vor allem beim Definieren von Ausnahmen von allgemein blockierten Kontakten und Apps wurde das Feedback als mangelhaft empfunden. Die Versuchspersonen zeigten sich verunsichert darüber, ob eine getroffene Einstellung den eigenen Vorstellungen/Zielen entsprach und ob sie gespeichert wurde.

3.2.2.3. Interpretation

Bei der Einstellung der (OFFTIME)-Dauer lag weiterhin ein Sichtbarkeits-Problem vor. Lediglich das Mapping zwischen „blauem Kreis“ und seiner Funktion konnte entscheidend verbessert werden. Die Funktionsweise, also das Drehen des Kreises im Uhrzeigersinn, blieb weiterhin schlecht zu erkennen. Da diese Form der Zeiteinstellung auch nicht in anderen Apps üblich ist, konnten die Versuchspersonen sich nicht auf Konventionen und übliche Standards beziehen, was das Entdecken der Funktionsweise vereinfacht hätte. Es gestaltete sich also weiterhin für die Versuchspersonen schwierig, sich ein vollständiges konzeptuelles Modell von der Funktionseinheit zur Zeiteinstellung zu verschaffen.

Die Inkonsistenzen vor allem in der Abfolge der Einstellung von Ausnahmen beim Blockieren von Kontakten/Apps sorgten dafür, dass die Versuchspersonen Schwierigkeiten dabei hatten, sich einheitliche und funktionierende mentale Modelle über die Funktionen der

App zu bilden. Die Inkonsistenz zwischen Anzeigetext und Bedienelement für den Zugang zum vertikalen Feineinstellungsmenü („Zahnräder“) sorgten für Frustration und Unsicherheit. Es wurden eine gute Sichtbarkeit und ein klares Mapping suggeriert, allerdings war beides tatsächlich nicht gegeben. Dies führte bei einigen Versuchspersonen dazu, dass sie sich selbst als Ursache des Problems sahen und der Meinung waren, sie würden die Funktionsweise der Menüs nicht verstehen.

Das stellenweise mangelnde Feedback führte außerdem zu Frustration. Der Status des Systems war bei einigen Einstellungen nicht klar nachvollziehbar, es herrschte Unsicherheit über die Folgen von Einstellungen. Die einzige Möglichkeit, Sicherheit über den Status der Einstellungen zu erhalten, war für die Versuchspersonen häufig, Schritte doppelt zu gehen und so zu überprüfen, ob die neuen Einstellungen im System wiedergefunden werden konnten.

3.2.2.4. Designimplikationen

Um die Anwendung der Zeiteinstellungsfunktion weiter zu verbessern, sollte versucht werden, die Art der Bewegung des Kreises bereits vor dessen Verwendung anzudeuten. Es sollte nicht nur die grundsätzliche Beweglichkeit des „blauen Kreises“ sichtbar gemacht werden, sondern auch, auf welchem Weg er zu bewegen ist. Dazu wurde eine Animation integriert, die jedes Mal beim Öffnen der App den „blauen Kreis“ eine Runde auf der Uhr „einfliegen“ lässt. So sollte das mentale Modell der Nutzer von einem reinen „Button“ hin zu einem interaktiven, beweglichen Bedienmodul gelenkt werden. Außerdem galt es, den Status der Zeiteinstellung noch klarer sichtbar zu machen. Die Nutzer sollten auf den ersten Blick erkennen können, welche Dauer, Start- und Endzeit zu jedem Zeitpunkt in der App eingestellt sind. Dafür sollte nun nicht nur die End-, sondern auch die Startzeit während der Zeiteinstellung eingeblendet werden. Außerdem wurde eine dauerhafte Anzeige der (OFFTIME)-Dauer integriert (s. Abb. 12). Somit sollte sichergestellt werden, dass durch die räumliche Verbindung auch ein inhaltliches Mapping erfolgt.

Als letztes sollte dafür gesorgt werden, dass die Grenzen der Zeiteinstellung besser deutlich gemacht werden. Um dies sicherzustellen wurde ein Pop-Up-Fenster erstellt, das bei einer Einstellung von mehr als 12 Stunden darauf verweist, dass hierfür die Nutzung der Kalenderintegration notwendig ist.

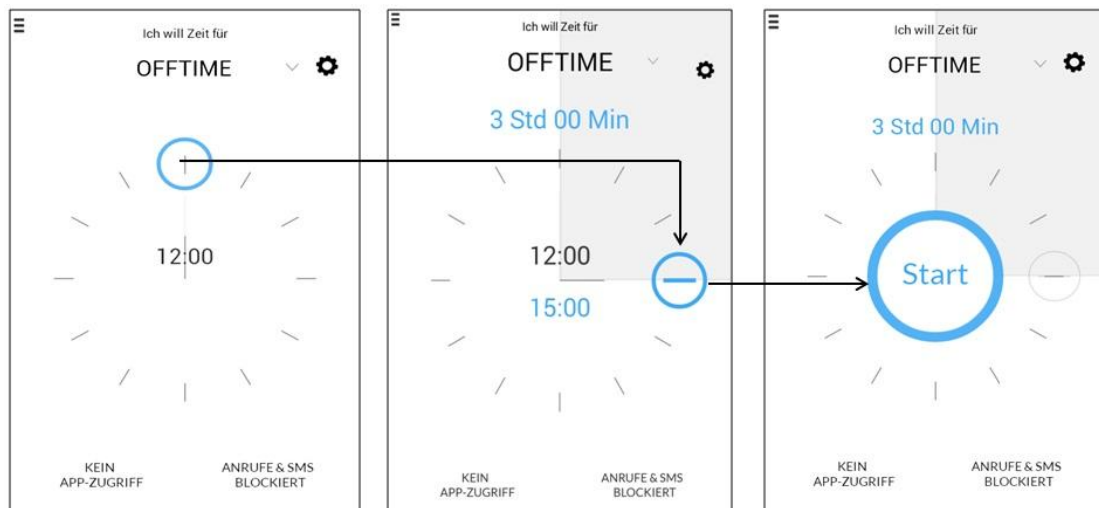


Abb. 12: Veränderte Funktionseinheit zur Einstellung der (OFFTIME)-Dauer nach Iteration 2

Um den beschriebenen Inkonsistenzen zu begegnen, wurde der letzte Screen des sequentiellen Menüs bearbeitet. Es wurde der Verweis auf die „Zange“ durch den Verweis auf die „Zahnräder“ ersetzt. Außerdem wurde beim Erstellen eines neuen Profils im letzten Screen des sequentiellen Menüs in der oberen rechten Bildschirmcke das „Zahnräder“-Symbol ergänzt. Bisher mussten Nutzer das sequentielle Menü erst verlassen und erneut betreten, um Zugang zum vertikalen Menü zu erhalten.

Desweiteren wurde die Struktur der Blockierung und Definition von Ausnahmen bei Kontakten und Apps angeglichen. Zu Beginn wurden nun häufigste Kontakte/Apps gefiltert und angezeigt. Bei einer Verwendung der Suchfunktion bezog sich die Suche allerdings stets auf alle Kontakte und Apps, unabhängig vom angezeigten Filter. Dadurch mussten Nutzer nicht wissen, in welchem Filter sich ein Kontakt oder eine App befindet. Das gesuchte Objekt konnte immer gefunden werden.

3.2.3. Dritter Testzyklus

Im dritten Testzyklus wurde ein Prototyp verwendet, in den die beschriebenen Designimplikationen aus Iteration zwei integriert worden waren. Die Testaufgaben blieben konstant.

3.2.3.1. Erwartungen

Nachdem die Veränderungen an der Funktionseinheit zum Einstellen der (OFFTIME)-Dauer und Starten eines Profils in der zweiten Iteration weitaus umfangreicher waren, wurde nun erwartet, dass die hier auftretenden Probleme behoben werden konnten.

Durch die Beseitigung der Inkonsistenzen zwischen Anzeigetext und Symbolik sollte der Zugang zum vertikalen Menü ebenfalls keine Probleme mehr bereiten.

Auch bei der Ausnahmeeinstellung für blockierte Kontakte und Apps wurden durch die Umstrukturierung der Anzeige- und Suchfunktion keine Schwierigkeiten mehr erwartet.

3.2.3.2. Ergebnisse

Die Zeiteinstellung über das Bedienelement „Blauer Kreis“ funktionierte im dritten Testzyklus bei vier der fünf Versuchspersonen ohne jegliche Probleme. Die Funktionsweise des „blauen Kreises“ wurde nun gut erkannt, die Dauer der (OFFTIME) konnte ohne Schwierigkeiten definiert werden. Die Begrenzung der Einstellungsmöglichkeiten, 12 Stunden, wurde ebenfalls gut erkannt und die Alternative der Kalenderintegration verstanden.

Bezüglich des Zugangs zu den Einstellungsmenüs blieb ein Problem weiterhin bestehen. Obwohl die Symbole „Zahnrad“/„Zahnräder“ korrekt als Menüzugänge identifiziert wurden, war den Versuchspersonen teilweise nicht klar, welche Einstellungen sich konkret in den Menüs verbergen. Die Möglichkeiten und Grenzen der beiden Menüformen mussten weiterhin über Versuch und Irrtum erforscht werden.

Die Ausnahme von Kontakten/Apps von der allgemeinen Blockierung bereitete nach der Umgestaltung und Anpassung der Anzeige- und Suchfunktionen keine Schwierigkeiten mehr.

Im Vergleich zu den ersten beiden Testzyklen zeigte sich ein deutlicher Zugewinn bei der Flüssigkeit im Bedienen der App. Ein Großteil der Aufgaben, vor allem zum Einstellen der Dauer und Verwenden des sequentiellen Menüs, konnte von fast allen Versuchspersonen ohne großes Zögern und Nachdenken gelöst werden. Außerdem wurde eine deutlich positivere Einstellung der Versuchspersonen während der Testungen beobachtet.

3.2.3.3. Interpretation

Eine wichtige Veränderung der Zeiteinstellungsfunktion war die Integration der Animation des „blauen Kreises“. Obwohl das „Einfliegen“ von keiner Versuchsperson bewusst wahrgenommen wurde, reichte dieser unbewusst verarbeitete Reiz offenbar als Trigger aus, um die Beweglichkeit und Funktionsweise des „blauen Kreises“ zu erkennen. Die veränderte Zeit- und Daueranzeige verbesserte die Sichtbarkeit des Status der Funktion deutlich. Die Versuchspersonen hatten keine Schwierigkeiten mehr beim Zuordnen der Start- und Endzeit, sowie der eingestellten Dauer der (OFFTIME). Das Pop-Up mit dem Hinweis auf die Kalenderintegration sorgte für ein besseres Verständnis der Grenzen der Zeiteinstellung. Außerdem wurden die Versuchspersonen auf die Feineinstellungen der App aufmerksam gemacht. Somit diente das Pop-Up nicht nur als Hinweis für die Lösung eines Fehlers/Problems, sondern weckte auch Neugierde bei den Versuchspersonen und den Willen, bisher unbekannte Einstellungsfunktionen zu entdecken. Auf diese Weise wurden sie auf den Weg gebracht, sich eingehender mit den Möglichkeiten der App zu beschäftigen.

Die Umgestaltung der Definition von Ausnahmen bei blockierten Kontakten/Apps ermöglichte es den Versuchspersonen, sich ein einheitliches konzeptuelles Modell der Funktionen zu bilden und dieses auf beide Fälle anzuwenden. Da kein Widerspruch zwischen beiden Funktionen mehr bestand, konnte das einmal gelernte Verhaltensmuster ohne Schwierigkeiten als Konvention weiter verwendet werden.

Im Allgemeinen entstand der Eindruck, dass die Versuchspersonen im dritten Testzyklus deutlich weniger Frustration und Verwirrung zeigten. Das Verständnis für die Funktionen der App und die Funktionsweise der Bedienelemente konnte verbessert werden. Die Toleranz der Versuchspersonen gegenüber Fehlern stieg spürbar. Durch die Kombination aus verbesserter Sichtbarkeit der Funktionen/des Systemstatus und vereinfachter Bildung eines konzeptuellen Modelles bezüglich des Bedienens der App, wurden Einschränkungen der Usability und eigene Fehler als weniger negativ bewertet. Somit konnte insgesamt die Zufriedenheit bei der Nutzung der App im Sinne der ISO 9241 - 11 gesteigert werden.

3.2.4. Iterationsübergreifende Ergebnisse

Zusätzlich zu den beschriebenen Usability-Schwachstellen konnten weitere Problemfelder identifiziert werden. Obwohl sie von den Versuchspersonen als unangenehm oder verwirrend

empfunden wurden, mussten sie aus Zeitgründen zurückgestellt werden oder wurden durch das Entwicklerteam als zweitrangige Probleme eingeordnet. Aus diesem Grund sind sie nicht weiter in die vorliegende Studie eingegangen und es wurden hierzu keine Designlösungen abgeleitet.

Zu diesen nicht weiter betrachteten Schwachstellen gehörten zum Beispiel einige Beschriftungen/Anzeigen, die über verschiedene Einstellungsstufen hinweg nicht konsistent waren. Außerdem wurde ein Wechsel zwischen englischer und deutscher Sprache beanstandet, und die starke Textlastigkeit der Screens des sequentiellen Menüs bemängelt. Unvollständige Textlabels erschwerten einige Eingaben. Pop-Up-Anzeigen, die den Nutzer etwa über gesperrte Apps informieren, erschienen so kurz, dass es kaum möglich war, sie vollständig zu lesen und die Informationen zu verstehen. Schwierigkeiten bereitete außerdem die Verwendung der Kalenderintegration.

3.3. Summative Evaluation

Im Anschluss an den iterativen Prozess der formativen Evaluation, wurde im dritten Schritt eine quantitative Auswertung der Daten durchgeführt.

Bei Aufgabe 1 wurde die Verwendung der Funktionseinheit zur Einstellung der Dauer eines Profils und dessen Start überprüft. Es zeigte sich im Laufe der Iterationen ein Trend zur Steigerung der Effizienz. Im ersten Testzyklus benötigten die Versuchspersonen im Mittel 2,75 (SD = 2,17), im zweiten Testzyklus 3 (SD = 3,55), und im dritten Testzyklus nur noch 1,12 (SD = 1,32) Klicks mehr als bei der optimalen Lösung. Auch die Effektivität der Verwendung der Funktionseinheit konnte gesteigert werden. In den ersten beiden Testzyklen lag die Lösungshäufigkeit bei 4, im dritten Testzyklus bei 5.

Aufgabe 4 beschäftigte sich erneut mit der Funktionseinheit zur Einstellung der Dauer eines (OFFTIME)-Profils. Es zeigte sich eine deutliche Steigerung der Effizienz von Iteration zu Iteration. Im ersten Testzyklus wurden durchschnittlich 16 (SD = 9), im zweiten Testzyklus 4 (SD = 4) und im dritten Testzyklus 0 (SD = 0) Klicks mehr als in der optimalen Lösung benötigt. Die Effektivität konnte ebenfalls gesteigert werden. Im ersten und zweiten Testzyklus wurde die Aufgabe jeweils von 2, im dritten Testzyklus von 4 Versuchspersonen korrekt gelöst.

Auch Aufgabe 6h bezog sich auf die Bedienung der Zeiteinstellungsfunktion. (1. Testzyklus: $M = 0,3$ ($SD = 0,5$) Klicks mehr; 2. Testzyklus: $M = 0,4$ ($SD = 0,5$) Klicks mehr; 3. Testzyklus: $M = 0,4$ ($SD = 0,8$) Klicks mehr; Lösungshäufigkeit 1. Testzyklus: 3; 2. und 3. Testzyklus: 5).

Die Entwicklung von Effizienz und Effektivität der Funktionseinheit zur Zeiteinstellung kann in Abb. 13 und Abb. 14 nachvollzogen werden.

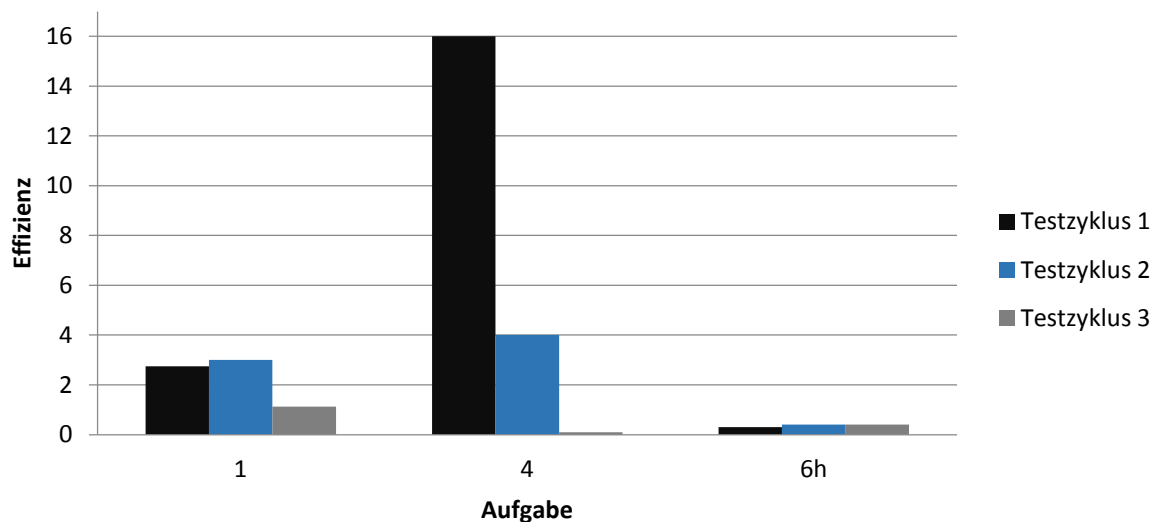


Abb. 13: Entwicklung der Effizienz der Funktionseinheit zur Einstellung der (OFFTIME)-Dauer

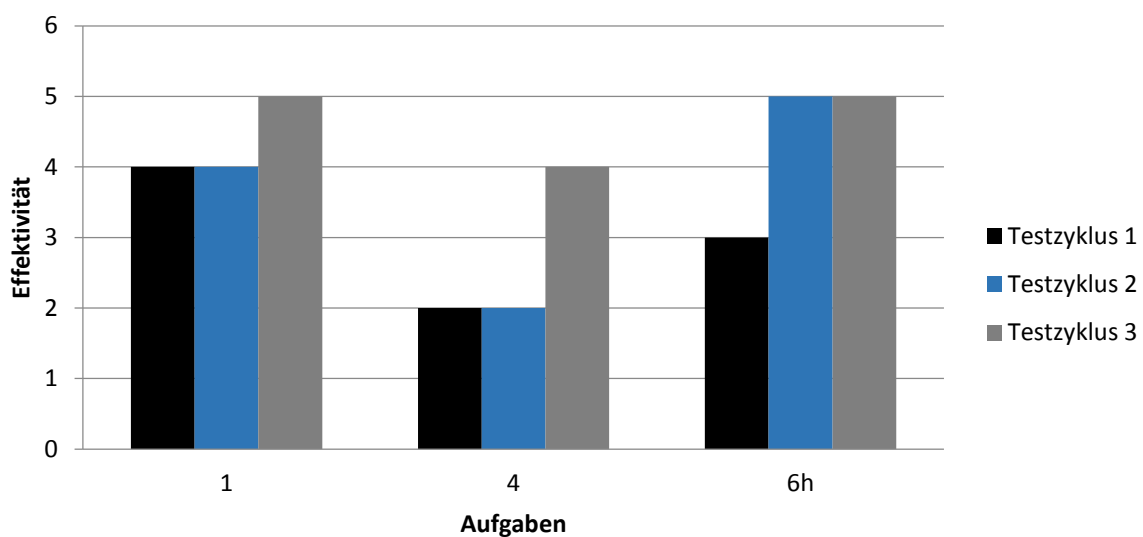


Abb. 14: Entwicklung der Effektivität der Funktionseinheit zur Einstellung der (OFFTIME)-Dauer

Aufgabe 6a zeigte eine verbesserte Effizienz bei der Funktion zum Hinzufügen eines Profils nach der ersten Iteration (1. Testzyklus: $M = 3,6$ ($SD = 4$) Klicks mehr; 2. Testzyklus: $M = 0,4$ ($SD = 0,8$) Klicks mehr; 3. Testzyklus: $M = 1$ ($SD = 1,3$) Klicks mehr). Die Effektivität war konstant durch eine Lösungshäufigkeit von 5 in allen Testzyklen.

Aufgabe 6b legte eine schrittweise Steigerung der Effizienz bei der Blockierung von Kontakten und Definition von Ausnahmen nahe (1. Testzyklus: $M = 11,7$ ($SD = 10,8$) Klicks mehr; 2. Testzyklus: $M = 2,8$ ($SD = 0,7$) Klicks mehr, 3. Testzyklus: $M = 1,2$ ($SD = 1,9$) Klicks mehr). Die Effektivität steigerte sich nach dem ersten Testzyklus, in dem eine Lösungshäufigkeit von 3 vorlag. Im zweiten und dritten Testzyklus lag sie bei 5.

In Aufgabe 6d sollte die gelernte Abfolge reproduziert werden (1. Testzyklus: $M = 0,6$ ($SD = 0,8$) Klicks mehr; 2. Testzyklus: $M = 0,2$ ($SD = 0,4$) Klicks mehr; 3. Testzyklus: $M = 0,8$ ($SD = 1,2$) Klicks mehr; Lösungshäufigkeit konstant 5). Die Ergebnisse sind in Abb. 15 veranschaulicht.

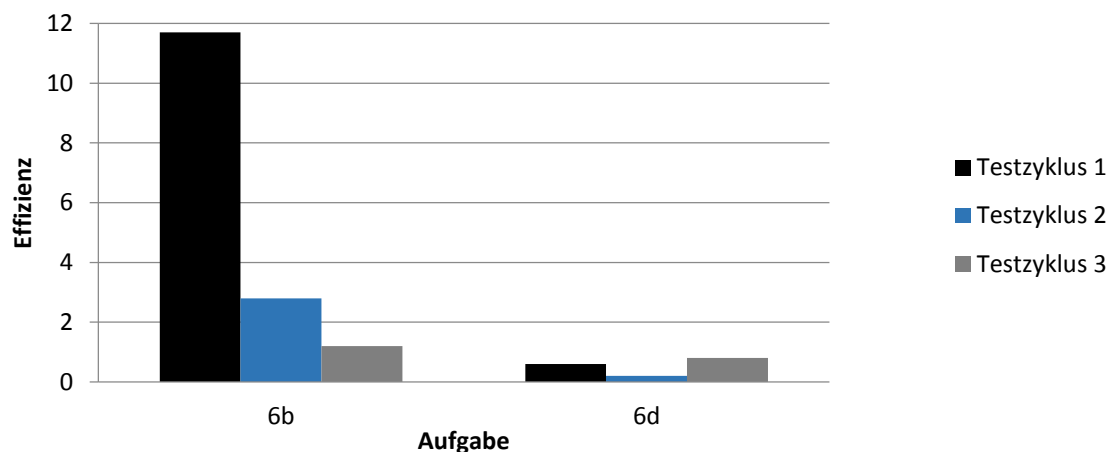


Abb. 15: Entwicklung der Effizienz der Funktionseinheit zur Blockierung von Kontakten/Apps und der Definition von Ausnahmen

In Aufgabe 6e zeigte sich, dass in der letzten Iteration eine Steigerung der Effizienz beim Eintritt in das vertikale Menü erreicht werden konnte (s. Abb. 16; 1. Testzyklus: $M = 2$ ($SD = 1,6$) Klicks mehr; 2. Testzyklus: $M = 9,1$ ($SD = 4,3$) Klicks mehr; 3. Testzyklus: $M = 0,8$ ($SD = 1,3$) Klicks mehr). Die Lösungshäufigkeit betrug im ersten Testzyklus 3, im zweiten Testzyklus 5 und im dritten Testzyklus 4.

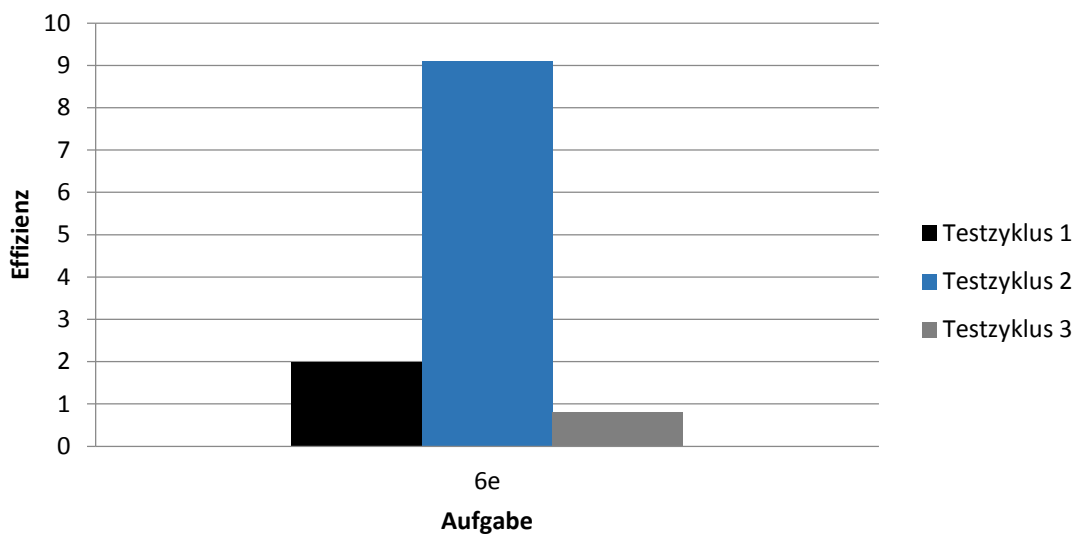


Abb. 16: Effizienz beim Eintritt in das vertikale Menü

Eine Tendenz zur Abnahme der Effizienz beim Wechsel eines Profils und Eintritt in das sequentielle Menü wurde in Aufgabe 5 beobachtet. Die Versuchspersonen benötigten im ersten Testzyklus durchschnittlich 3,6 (SD = 3), im zweiten Testzyklus 4,2 (SD = 2,1) und im dritten Testzyklus 8,2 (SD = 4,8) Klicks mehr als in der optimalen Lösung. Die durch die Aufgaben gemessene Effektivität war konstant. In allen Testzyklen lag die Lösungshäufigkeit bei 5.

Durch Aufgabe 6f wurde die Effizienz und Effektivität der Kalenderfunktion gemessen. Im ersten Testzyklus wurden hier durchschnittlich 4 (SD = 1,4), im 2. Testzyklus 2 (SD = 0,7), und im dritten Testzyklus 10,5 (SD = 2,5) Klicks mehr benötigt, als in der optimalen Lösung. Die Lösungshäufigkeit lag im ersten und dritten Testzyklus bei 2, im zweiten Testzyklus bei 4.

Bei Aufgabe 2 und 3 handelte es sich um einfache Zwischenaufgaben während der Laufzeit eines Profils und dessen Beendigung. Die hierfür gemessene Effizienz und Effektivität war annähernd konstant (Aufgabe 2: 1. Testzyklus: M = 0,2 (SD = 0,4) Klicks mehr; 2. Testzyklus: M = 0 (SD = 0) Klicks mehr; 3. Testzyklus: M = 0 (SD = 0) Klicks mehr; Aufgabe 3: 1. Testzyklus: M = 0,25 (SD = 0,43) Klicks mehr; 2. Testzyklus: M = 1 (SD = 1,26) Klicks mehr; 3. Testzyklus: M = 0,4 (SD = 0,8) Klicks mehr). Die Aufgaben konnten in allen Testzyklen von allen Versuchspersonen gelöst werden, die die Aufgabe

begonnen haben. Es gab je eine Person, die im ersten Testzyklus die 3. und im dritten Testzyklus die 2. Aufgabe nicht begonnen hat.

Bei Aufgabe 6c handelte es sich erneut um eine einfache Zwischenaufgabe. Die Versuchspersonen erstellten eine automatische Antwort-SMS. Hier konnten ebenfalls eine stabile Effizienz und Effektivität über die Iterationen hinweg festgestellt werden (Aufgabe 6c: 1. Testzyklus: $M = 1,6$ ($SD = 1,9$) Klicks mehr; 2. Testzyklus: $M = 1,2$ ($SD = 1,6$) Klicks mehr; 3. Testzyklus: $M = 1,6$ ($SD = 2,3$) Klicks mehr; Lösungshäufigkeit konstant 5).

Aufgabe 6g war abermals eine einfache Zwischenaufgabe (1. Testzyklus: $M = 3,2$ ($SD = 4,1$) Klicks mehr; 2. Testzyklus: $M = 3,2$ ($SD = 4,7$) Klicks mehr; 3. Testzyklus: $M = 0$ ($SD = 0$) Klicks mehr; Lösungshäufigkeit 1. und 2. Testzyklus: 5; 3. Testzyklus: 4).

4. Diskussion

Durch die Kombination aus heuristischer und formativer Evaluation konnten im Laufe der drei Iterationen eine Vielzahl an Usability-Problemen in der App (OFFTIME) identifiziert werden. Die Design-Schwachstellen, die in der heuristischen Evaluation ausfindig gemacht wurden, bestätigten sich in der formativen Evaluation. Allerdings konnten die Probleme dank der qualitativen Daten aus den Think-Aloud Protokollen wesentlich spezifischer und besser aus Sicht der Nutzer betrachtet und beschrieben werden. Erst diese genaue Untersuchung der verschiedenen Facetten der Schwierigkeiten, die die einzelnen Versuchspersonen mit den Funktionseinheiten der App hatten, machte es möglich, präzise Designimplikationen abzuleiten. Dass diese ihr Ziel, die Usability der App zu verbessern, in weiten Teilen erreichen konnten, zeigte sich durch die streckenweise sehr flüssige Nutzung der App durch die Versuchspersonen im dritten Testzyklus. Vor allem die Einstellung der (OFFTIME)-Dauer, der Zugang zu den Einstellungsmenüs und die Funktion zum Hinzufügen eines Profils wurden im letzten Testzyklus deutlich besser verstanden. Subjektive Erfolgserlebnisse beim korrekten Vornehmen von Einstellungen waren häufiger zu beobachten, die Versuchspersonen zeigten sich gegenüber eigenen Fehlern toleranter und neigten weniger zu Frustration. Unterm Strich wiesen die Ergebnisse der quantitativen Analyse auf eine deutliche Steigerung der Zufriedenheit der Versuchspersonen beim Gebrauch der App nach Ende der drei Iterationen hin.

Nach der ISO 9241-11 (1999) wird Usability allerdings neben der Zufriedenheit der Nutzer, auch durch die Aspekte Effektivität und Effizienz beschrieben. Bei den

Funktionseinheiten zum Einstellen der (OFFTIME)-Dauer, Hinzufügen eines Profils, sowie Blockieren von Kontakten/Apps und Definieren von Ausnahmen konnte in diesen Bereichen eine teilweise deutliche Steigerung erreicht werden. Der Eintritt in das vertikale Menü konnte, nachdem die Probleme der Sichtbarkeit, des Mappings und der Inkonsistenz beseitigt wurden, im Prototypen für die dritte Iteration ebenfalls effizienter gestaltet werden.

Ein Widerspruch zwischen den Ergebnissen der quantitativen und qualitativen Analyse zeigte sich bei der Betrachtung des Eintritts in das sequentielle Menü. Bei der formativen Evaluation wurde deutlich, dass das Verständnis für das Symbol „Zahnrad“ gegenüber dem „Zauberstab“ deutlich größer war. Außerdem hatten die Versuchspersonen ein klareres mentales konzeptuelles Modell über die Funktionen, die sich hinter dem Symbol verbergen könnten. Bei der summativen Evaluation allerdings zeigte sich eine tendenzielle Abnahme der Effizienz dieser Funktion.

Es gibt zwei primäre Limitationen der Studie, die zu dieser Diskrepanz geführt haben könnten. Es muss kritisch betrachtet werden, dass die Aufgaben häufig nicht eindimensional formuliert waren. In Aufgabe 5 etwa wurde das erste Mal das sequentielle Menü für die Lösung benötigt. Allerdings musste erst ein Profilwechsel vorgenommen werden, bevor im sequentiellen Menü der Profilname geändert werden konnte. Es kann also nicht mit Sicherheit festgestellt werden, ob die Schwierigkeiten vielleicht eher im Profilwechsel selbst lagen.

Der vermutlich noch wichtigere Aspekt ist der, dass die Stichproben von 5 Versuchspersonen für eine klassische quantitative Analyse deutlich zu klein sind. Es kann nicht eindeutig davon ausgegangen werden, dass sie die Werte der Grundgesamtheiten repräsentativ vorhersagen. Ein statistischer Vergleich von Mittelwerten, die aus lediglich 5 Einzelwerten gebildet wurden und darüber hinaus nicht konstant über homogene Varianzen verfügen (Bortz & Schuster, 2010), ist im besten Fall kritisch zu betrachten. Somit kann keine klare Aussage über die statistische Bedeutsamkeit der Unterschiede zwischen den Mittelwerten der Testzyklen getroffen werden. Die Ergebnisse der summativen Evaluation sind also lediglich als Trends zu interpretieren. Die Wahl der Stichprobengröße ist allerdings durch den Fokus der Studie, der auf der formativen Evaluation lag, zu begründen. Hierbei wird eine Stichprobengröße von 5 Versuchspersonen je Testzyklus als geeignetes Verhältnis zwischen Kosten und Nutzen betrachtet (Nielsen, 2000).

Eine weitere Limitation der Studie war, dass nicht allen identifizierten Usability-Problemen mit Veränderungen im Design begegnet wurde. Dies lag vornehmlich am begrenzten Zeitkontingent für die Iterationen, aber auch an der Priorisierung durch die

(OFFTIME)-Entwickler. Es konnte somit nicht untersucht werden, inwieweit einige Problemfelder eventuell miteinander verknüpft sind. So ist zum Beispiel denkbar, dass sich die Nutzung der Kalenderintegration und Einstellungsmenüs bereits durch ein allgemein verbessertes Feedback vereinfachen würden – auch ohne, dass diese Funktionen inhaltlich und strukturell verändert werden müssten.

Zusammenfassend kann festgehalten werden, dass durch das iterative, entwicklungsbegleitende Vorgehen einige der größten Mängel des Designs in Hinblick auf die Usability der (OFFTIME)-App beseitigt werden konnten. Es gilt aber vor allem bei der Kalenderintegration und der generellen Einarbeitung von Feedback, das Design der App weiter zu verfeinern.

Ein möglicher nächster Schritt auf dem Weg zu einer weiter gesteigerten Usability der App könnte sich auch auf die Überarbeitung der Menüstrukturen beziehen. Die Undurchsichtigkeit der Einstellungsoptionen – welche Einstellungen sind in welchem Menü vorzunehmen? – sorgte bei vielen Versuchspersonen für große Unsicherheit. Fraglich ist, ob eine vereinheitlichte Struktur des Einstellungsmenüs anstelle der zwei getrennten Strukturen nicht einfacher zu verstehen und zu nutzen wäre. Hierin könnte ein Ansatz für eine weiterführende Untersuchung der App liegen.

In diesem Zusammenhang wäre es zudem möglich, die generellen Vor- und Nachteile unterschiedlicher Menüformen zu betrachten. Es sollte durch die steigende Bedeutung von interaktiven technischen Geräten, zu denen auch Smartphones gehören, ein wichtiges Forschungsziel sein, Menüstrukturen zu identifizieren, die verschiedenen Ansprüchen gerecht werden können. Es wird nach Lösungen gesucht, die eine maximale Performanz des Nutzers ermöglichen und außerdem eine hohe Zufriedenheit bei ihrer Nutzung erzeugen können (Yu & Roh, 2002; Leuthold, Schmutz, Bargas-Avila, Tuch & Opwis, 2011). Außerdem sollen die Menüs an die unterschiedlichen Bedürfnisse der Nutzer oder verschiedene Aufgabentypen angepasst werden (Bernard & Hamblin, 2003). Ein dritter Ansatz ist, die Menüform entsprechend der Eigenschaften unterschiedlicher Nutzergruppen zu gestalten (Ziefle & Bay, 2004). Allerdings beziehen sich diese Forschungsansätze bisher vornehmlich auf das Design von Webseiten. Ob die Erkenntnisse eins zu eins auf das Design von Smartphone-Apps übertragen werden können, ist fraglich. Das Problem liegt hierbei weniger in der Unterscheidung Website versus App. Vielmehr müssen die verschiedenen Ansprüche und technischen Möglichkeiten differenziert werden, die beim Gebrauch von größeren Desktop-Geräten und portablen Mobile-Devices zum Tragen kommen.

Im Bereich des User-Interface- und User-Experience-Design wird heute häufig die Methode des Mobile-First-Designs angewendet. Hierbei wird zunächst das Design für mobile Geräte entwickelt, bevor die Anpassung für klassische Geräte wie Desktop-PC und Laptop erfolgt. Die Annahme ist, dass die Entwicklung des sehr reduzierten, sparsamen und fokussierten Designs für die kleinen portablen Geräte als gute Grundlage für die spätere Entwicklung eines schrittweise erweiterten und ergänzten Designs für die größeren Bildschirme der Desktop-Geräte dient (Johnson, 2013; Camenisch, 2014). Wenn allerdings gar nicht genau geklärt ist, mit welchen Prinzipien und Strukturen ebendieses sparsame Design optimal implementiert werden kann, gehen die Vorteile eines speziellen Designs für Mobile-Devices unter Umständen verloren.

Das Ziel, das alle diese Überlegungen gemein haben, ist, das Design von Webseiten oder Anwendungen – wie (OFFTIME) – zu optimieren und dadurch ihre Usability zu steigern. Nur dann können die Nutzer zufriedengestellt und an das jeweilige Produkt gebunden werden. Im Fall der (OFFTIME)-App ist mit diesem Aspekt auch die eigentliche inhaltliche Frage- und Zielstellung der Entwickler von (OFFTIME) eng verbunden. Kann die App tatsächlich durch eine Regelung der Smartphone-Nutzung zu einer Entlastung im täglichen Leben führen? Und sind das Design und die Struktur von (OFFTIME) so gestaltet, dass sie diesen Zweck aktiv unterstützen, anstatt ihm entgegenzuwirken? Denn nur, wenn das Design übersichtlich und verständlich, einfach zu bedienen, gut strukturiert, dabei ansprechend und interessant gestaltet ist, können die bedienungsbezogenen Fragestellungen in den Hintergrund, die arbeits-, konzentrations- und entspannungsbezogenen Gewinne in den Vordergrund treten.

Die vorliegende Studie konnte auf diesem Weg erste Schritte gehen. Dennoch gilt es natürlich weiterhin, (OFFTIME) kontinuierlich weiterzuentwickeln und den Bedürfnissen der Nutzer noch präziser anzupassen.

Literaturverzeichnis

- Bernard, M. & Hamblin, C. (2003). Cascading versus indexed menu design. *Usability News*, 5 (1). Zugriff am 20.08.2014. Verfügbar unter <http://usabilitynews.org/cascaded-versus-indexed-menu-design/>
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Auflage). Berlin, Heidelberg: Springer-Verlag.
- Camenisch, J. (2014). *7 Things You Need To Know About Mobile Web Design*. Zugriff am 24.08.2014 unter <https://www.odesk.com/blog/2014/04/7-things-need-know-mobile-web-design/>
- Elis, A. (2011). *Kommunikations-Stress: Die Smartphone-Sklaven*. Zugriff am 17.08.2014 unter <http://www.spiegel.de/politik/deutschland/kommunikations-stress-die-smartphone-sklaven-a-789420.html>
- Grauers, J. & Wall, E. (2012). *Preventing stress through smartphone usage*. Unveröffentlichte Bachelorarbeit, Insitut für Angewandte Informatik, Universität Göteborg.
- Greenhouse, E. S (n.d.). *Human-Centered Design*. Zugriff am 10.08.2014 unter <http://www.aging.ny.gov/livableny/resourcemanual/demographicandsocialtrends/i9.pdf>
- Gould, J.D. (1988). How to Design Usable Systems. In: M. Helander (Ed.), *Handbook of Human Computer Interaction* (pp. 757–789). Amsterdam: Elsevier Science Publishers B.V.

- Hartson, H. R., Andre, T. S. & Williges, R. C. (2001). Criteria for evaluating usability evaluation methods. *International journal of human-computer interaction*, 13 (4), 373-410.
- Hegner, M. (2003). *Methoden zur Evaluation von Software* (IZ-Arbeitsbericht Nr. 29). Bonn: InformationsZentrum Sozialwissenschaften.
- ISO 9241-11 (1999). *Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten. Teil 11: Anforderungen an die Gebrauchstauglichkeit – Leitsätze*. Berlin: Beuth.
- Johnson, J. (2013). *Mobile First Design: Why It's Great and Why It Sucks*. Zugriff am 24.08.2014 unter <http://designshack.net/articles/css/mobilefirst/>
- Kaspar, K., Hamborg, K. C., Sackmann, T. & Hesselmann, J. (2010). Die Effektivität formativer Evaluation bei der Entwicklung gebrauchstauglicher Software – eine Fallstudie. *Zeitschrift für Arbeits-und Organisationspsychologie A&O*, 54 (1), 29-38.
- Krug, S. (2010). *Web Usability – Rocket Surgery made easy*. München: Addison-Wesley Verlag.
- Leuthold, S., Schmutz, P., Bargas-Avila, J. A., Tuch, A. N. & Opwis, K. (2011). Vertical versus dynamic menus on the world wide web: Eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. *Computers in Human Behavior*, 27 (1), 459-472.
- Lewis, J.R. (2014). Usability: Lessons Learned ... and Yet to Be Learned. *International Journal of Human-Computer Interaction*, 30 (9), 663-684.

- Lewis, C., Polson, P. G., Wharton, C. & Rieman, J. (1990). Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. *Proceedings of the CHI '90 Conference on Human Factors in Computing Systems*, 235-242. New York: ACM.
- Nielsen, J. (1992). The usability engineering life cycle. *Computer*, 25 (3), 12-22.
- Nielsen, J. (1993). *Usability engineering*. Boston: Academic Press.
- Nielsen, J. (1995). *10 Usability Heuristics for User Interface Design*. Zugriff am 20.04.2014 unter <http://www.nngroup.com/articles/ten-usability-heuristics/>
- Nielsen, J. (2000). *Why you only need to test with 5 Users*. Zugriff am 30.03.2014 unter <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- Nielsen, J. (2011). *Parallel & Iterative Design + Competitive Testing = High Usability*. Zugriff am 18.02.2014 unter <http://www.nngroup.com/articles/parallel-and-iterative-design/>
- Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.
- Rauterberg, M. (1991). Partizipative Konzepte, Methoden und Techniken zur Optimierung der Softwareentwicklung. *Softwaretechnik-Trends 11* (3). 104-126.
- Rieman, J., Franzke, M. & Redmiles, D. (1995). Usability evaluation with the cognitive walkthrough. *Proceedings of the CHI '95 Conference on Human Factors in Computing Systems*, 387-388. Denver: ACM.
- Shackel, B. (1991). Usability - Context, Framework, Definition, Design and Evaluation. In B. Shackel & S. Richardson (Eds.). *Human Factors for Informatics Usability* (pp. 21-37), Cambridge: Cambridge University Press.

- In Springer Gabler Verlag (Hrsg.), Gabler Wirtschaftslexikon (2011). Stichwort: *Smartphone*. Zugriff am 11.08.2014 unter <http://wirtschaftslexikon.gabler.de/Archiv/569824/smartphone-v1.html>
- Statista (2014). *Anzahl der Smartphone-Nutzer in Deutschland in den Jahren 2009 bis 2014 (in Millionen)*. Zugriff am 13.08.2014 unter <http://de.statista.com/statistik/daten/studie/198959/umfrage/anzahl-der-smartphonenuutzer-in-deutschland-seit-2010/>
- Statista (2014). *Anzahl der Smartphone-Nutzer in USA in den Jahren 2010 bis 2014 (in Millionen)*. Zugriff am 13.08.2014 unter <http://de.statista.com/statistik/daten/studie/285590/umfrage/anzahl-der-smartphone-nutzer-in-usa-bis-2013/>
- Thomée, S. (2012). *ICT use and mental health in young adults – Effects of computer and mobile phone use on stress, sleep disturbances, and symptoms of depression*. Unveröffentlichte Dissertation, Institut für Medizin der Sahlgrenska Akademie, Universität Göteborg.
- Yu, B. M. & Roh, S. Z. (2002). The effects of menu design on information-seeking performance and user's attitude on the World Wide Web. *Journal of the American Society for Information Science and Technology*, 53 (11), 923-933.
- Yun, H., Kettinger, W. J. & Lee, C. C. (2012). A new open door: The smartphone's impact on work-to-life conflict, stress, and resistance. *International Journal of Electronic Commerce*, 16 (4), 121-152.
- Ziefle, M. & Bay, S. (2004). Mental models of a cellular phone menu. Comparing older and younger novice users. In S. Brewster & M. Dunlop (Eds.) *Mobile Human-Computer Interaction-MobileHCI 2004* (pp. 25-37). Berlin, Heidelberg: Springer-Verlag.

Anhang

1) Aufruf für Versuchspersonen

Für eine Bachelorarbeit am Lehrstuhl für Ingenieurpsychologie unter der Leitung von Prof. Hartmut Wandke und in Zusammenarbeit mit dem Start-Up Unternehmen (OFFTIME) werden Personen für eine Studie mit der neuen Smartphone-App (OFFTIME) gesucht.

Fast jeder kennt diese Situationen: Während des wichtigen Meetings klingelt plötzlich das Handy. Und obwohl man sich eigentlich auf die Vorlesung konzentrieren wollte, lockt das neue Spiel auf dem Smartphone. Diese Ablenkungen im Alltag gezielt abzuwehren und die Nutzung des Smartphones genauer an die aktuellen Bedürfnisse anzupassen, hat sich (OFFTIME) zur Aufgabe gemacht.

Doch, wie einfach lässt sich die App nutzen? Genau das soll das Thema der Studie sein! Nicht Ihre Fähigkeiten im Umgang mit Smartphones, sondern die Nutzerfreundlichkeit der App stehen hier im Fokus.

Wenn Sie Lust haben, die App (OFFTIME) in einem Test (Dauer ca. 45 min) kennenzulernen und ihre Funktionen zu erkunden, und wenn Sie

- ihr Smartphone *intensiv* nutzen,
- ihr Smartphone mit *Android*-Software ausgestattet ist und
- Sie noch *keine* Erfahrungen mit der App (OFFTIME) haben

sind Sie für uns genau der/ die Richtige!

Herzlich laden wir Sie ein, an einem der folgenden Termine an der Studie teilzunehmen.

Gern können auch individuell andere Termine vereinbart werden!!!

Alle Termine finden in Adlershof, Rudower Chaussee 18, Raum 1'229 statt.

Für Studierende der Psychologie winken 3 VP-Marken.

Wir freuen uns auf Ihre Anmeldung (unter jette.beisser@googlemail.com),

Jette Beißer & das (OFFTIME)-Team

2) Skript Einweisung Testprozedur

Fast jeder kennt Momente, in denen er oder sie gern einmal seine Ruhe haben möchte. Sei es beim Spieleabend mit der Familie, dem wichtigen Meeting auf Arbeit, oder während man gerade seinem liebsten Hobby nachgeht. Anrufe, E-Mails und verlockende Smartphone-Apps stören hier oft einfach. Doch das Handy komplett ausschalten, wer will das schon?

(*OFFTIME*) ist eine App, die es sich zur Aufgabe gemacht hat, ihren Nutzern den reduzierten Umgang mit dem Smartphone zu erleichtern. Sie soll in Situationen helfen, in denen die Nutzer nicht selbst auf bestimmte Kontakte reagieren, sich für eine reduzierte Nutzung von Apps selbst disziplinieren, oder einfach für eine gewisse Zeit ungestört sein wollen.

Hierfür wurde unter anderem die Möglichkeit eingerichtet, eingehende Telefonanrufe zu blockieren oder automatisch zu beantworten, Apps und Benachrichtigungen einzeln oder in ihrer Gesamtheit zu blockieren und den Internetzugang zu sperren.

In dieser Studie soll es nun darum gehen, herauszufinden, wie nutzerfreundlich die (*OFFTIME*)-App gestaltet ist. Um dies zu untersuchen werden Sie, liebe Studienteilnehmerin/ lieber Studienteilnehmer, darum gebeten, 5 Aufgaben mit der (*OFFTIME*)-App zu lösen.

Dabei wird nicht getestet, wie gut Ihre individuellen Fähigkeiten im Umgang mit Smartphones sind: Es steht einzig die App auf dem Prüfstand!

Jede Aufgabe wird dabei in 5 *Schritten* bearbeitet. Hierbei werden Sie darum gebeten, ihre Gedanken laut zu äußern und möglichst genau zu beschreiben.

Wir werden dieses Prozedere gleich gemeinsam an einer bekannten Anwendung (Kalender) üben:

- 1 Was ist auf dem Bildschirm zusehen? Welche Bedienelemente und Anzeigen sind insgesamt (unabhängig von der konkreten Aufgabe) zu erkennen? Bitte nennen Sie alles, was Sie wahrnehmen können! Auch, wenn Sie nicht genau wissen, was es ist und wozu es dient.
- 2 Nun geht es an die aktuelle Aufgabe:
Was denken Sie, welche Bedienelemente vermutlich betätigt werden müssen, um Aktionen auszulösen, die Sie dem Ziel dieser speziellen Aufgabe näher bringen?

- 3 Was wird passieren, wenn Sie das Bedienelement, z.B. eine Schaltfläche, betätigt haben? Welche konkreten Reaktionen der (OFFTIME)-App werden auf die Betätigung der Bedienelemente eintreten?

Bitte formulieren Sie ihre Vermutungen BEVOR sie die Handlung auslösen!

- 4 Ausführung der geplanten Handlung
- 5 Vergleich der erwarteten und tatsächlich beobachteten Reaktionen der Anwendung

Bitte lösen Sie die Aufgaben spontan und ohne Unterstützung von Hilfsmitteln (Google, Tutorials etc.)

3) Aufgaben

Übungsaufgabe:

Richten Sie einen einmaligen, 60-minütigen Termin am heutigen Tag um 20Uhr ein, der den Namen „Test“ trägt.

Aufgabe 1:

Starten Sie eine 60-minütige (OFFTIME). Verändern Sie die dafür getroffenen Profil-Voreinstellungen z.B. über App-Zugriffe und blockierte Kontakte **nicht**.

Aufgabe 2:

In dem Profil (OFFTIME) ist eingestellt, dass der Zugriff auf alle Apps gesperrt ist. Versuchen Sie dennoch, auf die App „Uhr“ zuzugreifen.

Aufgabe 3:

Beenden Sie die soeben gestartete (OFFTIME) vorzeitig und kehren Sie auf den Startbildschirm der App zurück.

Aufgabe 4:

Versuchen Sie, eine (OFFTIME)-Dauer von 13 Stunden einzustellen.

Aufgabe 5:

Die folgende Aufgabe besteht aus 3 Unterpunkten, die Sie bitte schrittweise abarbeiten:

- a) Wechseln Sie in das (OFFTIME)-Profil „Familie“
- b) Benennen Sie das Profil um in „Zu Hause“
- c) Kehren Sie zum Startbildschirm zurück

Aufgabe 6:

Nun simulieren wir die umfangreiche und vollständige Nutzung der (OFFTIME)-App. Dafür richten Sie nun selbst ein Profil ein und stellen spezielle Zugriffs- und Nutzungsrechte ein:

- a) Richten Sie ein neues Profil mit dem Namen „Uni“ ein
- b) Sorgen Sie dafür, dass alle Kontakte außer „Mama“ blockiert werden.
- c) Erstellen Sie eine automatische Antwort auf alle eingehenden blockierten Anrufe:
„Ich bin gerade in der Uni und ab %endtime wieder erreichbar.“
- d) Sperren Sie den Zugriff auf alle Apps außer „Chrome“
- e) Lassen sie den Internetzugang aktiv
- f) Stellen Sie nun ein, dass das Profil „Uni“ automatisch aktiv wird, wenn die im Google-Kalender gespeicherten Termine „Vorlesung“, „Übung“, und „Seminar“ stattfinden.
- g) Wählen Sie zur Beendigung der (OFFTIME) die Option „Lange drücken“ aus
- h) Starten Sie eine 90 Minütige „Uni“-(OFFTIME)

Bitte arbeiten Sie diese Punkte in der angegebenen Reihenfolge ab.

4) Übersicht Problementwicklung

Testreihe 1	Identifiziertes Usability-Problem	Testreihe 2	Problementwicklung	Testreihe 3	Problementwicklung
Aufgabe		Implikation		Implikation	
Einstellung der (OFFTIME)-Dauer über blauen Kreis	Problem der Sichtbarkeit -> Funktion des blauen Kreises nicht erkannt	Stilisierte Uhr von Beginn an	Problem teilweise gelöst: Funktion zur Zeiteinstellung besser erkannt, konkrete Bedienung/ Funktionsweise aber weiterhin nicht intuitiv	Animation, blauer Kreis „fliegt“ auf der Uhr ein, bewegt sich im Uhrzeigersinn und deutet Beweglichkeit an	Gelöst, bis auf eine VP (die einige Zwischenschritte brauchte) haben alle Personen sofort richtig gehandelt -> obwohl die Bewegung des Kreises nicht bewusst wahrgenommen wurde
Einstellung der (OFFTIME)-Dauer	Problem der Sichtbarkeit -> Dauer oder Startzeit eingestellt? Zeitangabe verschwindet, wenn blauer Kreis losgelassen wird	Zeitangabe bleibt stehen	Problem besteht weiterhin; keine Transparenz hinsichtlich der Start- und Endzeit	Start und Endzeit angezeigt, Startzeit fix, Endzeit passt sich bei Bewegung des Kreises an	Gelöst
13 Stunden Dauer einstellen (Grenzen der Zeiteinstellung erkennen)	Grenzen der Einstellungsmöglichkeiten werden nicht sichtbar	Keine Veränderung durchgeführt	Problem besteht weiterhin	Bei Überquerung der 12 Stunden -> Pop-up-Information, Verwendung der Kalenderintegration	Gelöst, Grenzen werden sichtbar, gleichzeitiger Hinweis auf die Möglichkeit weiterer Einstellungen im vertikalen Menü
Betretten des sequentiellen, bzw. vertikalen Menüs zur Konfiguration der (OFFTIME)	Mapping-Problem -> Zauberstab-Symbol erlaubt keine Rückschlüsse auf Funktion des Buttons; Keine (hierarchische) Verbindung zwischen Zauberstab & Zangen-Symbol, kein Mapping der Hierarchie von groben zu Feineinstellungen	Verwendung des Android-Zahnrad-Symbols für sequentielles Menü -> Mapping zu Einstellungen; Zahnräder für vertikales Menü als Symbol für Feineinstellungen -> Mapping der Hierarchie	Neues Problem entstanden -> Erkennung der Symbole funktioniert gut; aber: Inkonsistenzen bei Beschriftung auf letzter Seite des sequentiellen Menüs -> Zugang zu vertikalem Menü gestaltet sich schwierig	Inkonsistenz in Beschriftung/ Symbolik beseitigt	Teilweise gelöst: Beide Symbole als Zugänge zu Menüs erkannt, allerdings bereitet inhaltliche Gestaltung der Menüs Probleme: welche Aktionen/ Einstellungen lassen die einzelnen Menüs zu? Intuitiv nicht erkennbar, lediglich über trial & error herauszufinden
Hinzufügen eines neuen Profils	Räumliche Trennung von Profilwahl/ Profilerstellung bereitet Probleme -> semantisch verknüpfte Funktionen auch räumlich verknüpft vermutet	Zusatzfunktion im Drop-down Menü für Profilwahl (Profil hinzufügen)	Gelöst; doppelten Weg für Funktion aber beibehalten (Im Haupt- & Profilwahl-Menü)		
		Aufgabe	Identifiziertes Usability-Problem	Implikation	Problementwicklung
		Blockierung von Kontakten/ Apps, dann Ausnahmen definieren	Inkonsistenz im Ablauf der Einstellung von Ausnahmen, sowie undurchsichtige Suchfunktion der App	Angleichung der Anzeigen und Umgestaltung der Suchfunktion (in allen Kontakten/ Apps)	Gelöst

5) Übersicht Probleme ohne Umsetzung von abgeleiteten Designimplikationen

Aufgabe	Identifiziertes Usability-Problem	Mögliche Implikation	Mögliches Resultat
Blockierung von Apps/Kontakten Einstellen, dann Ausnahmen definieren	Mangelndes Feedback über Auswahl (Profilbild, häufig nicht verwendet -> Nutzer sieht wenn überhaupt nur „jemand“ ausgewählt, aber nicht wer; tw. Standard-Bild auch nicht als Auswahlhinweis erkannt), bei keiner Auswahl überhaupt kein Feedback	Zusätzlich zu Profilbild Kontaktname anzeigen; wenn keine Ausnahmen eingestellt, kurze Notiz auf Einstellungsbildschirm im sequentiellen Menü	Bessere Sichtbarkeit des Status der Einstellung
Kalender-Automatisierung	Funktionalität der Automatisierung schwer verständlich (Kalender auswählen -> alle Termine des Kalenders starten (OFFTIME), Stichworte wählen einzelne Termine innerhalb des Kalenders aus)	Gezielte Auswahl einzelner Termine durch Klick zulassen; zur Anzeige Kalenderansicht verwenden, auf Verwendung der dreigeteilten Einstellung verzichten	Auf bekanntes Format des Kalenders zurückgreifen, bekannte Konzepte nutzen

Selbstständigkeitserklärung

Hiermit erkläre ich, dass die vorliegende Bachelorarbeit

- eigenständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt wurde;
- erstmalig zu diesem Studiengebiet eingereicht wird;
- unter Beachtung und Kenntnis der Prüfungsordnung (Amtliches Mitteilungsblatt der HU Berlin Nr. 39/2009) erstellt wurde.

Ort, Datum

Unterschrift