



ulm university universität
uulm

Universität Ulm | 89069 Ulm | Germany

Fakultät für
Ingenieurwissenschaften,
Informatik und
Psychologie
Institut für Datenbanken
und Informationssysteme

Evaluierung einer mobilen Anwendung zur flexiblen Datenerhebung anhand einer Benutzerstudie

Bachelorarbeit an der Universität Ulm

Vorgelegt von:

Jasmina Bandic

jasmina.bandic@uni-ulm.de

Gutachter:

Prof. Dr. Manfred Reichert

Betreuer:

Johannes Schobel

2017

Fassung 13.03.2017

© 2017 Jasmina Bandic

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/de/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

Kurzfassung

Mobile Applikationen finden heutzutage in jedem Bereich unseres Lebens Verwendung. Folglich ist der Besitz eines mobilen Endgeräts selbstverständlich. Diese zwei Faktoren lassen sich ausnutzen, um zum Beispiel Daten im medizinischen oder psychologischen Bereich zu erheben.

Das Projekt QuestionSys wurde mit dem Ziel gestartet solche Datenerhebungen zu ermöglichen. Im Rahmen des Projekts wurde eine mobile Applikation entwickelt, die es möglich macht, im Konfigurator erstellte Fragebögen ganz einfach auf das mobile Endgerät zu exportieren und zu benutzen.

Die Applikation muss einige grundsätzliche Anforderungen erfüllen. Eine davon ist die Benutzbarkeit (engl. Usability). Die Applikation muss von Nutzern angenommen und benutzt werden. Damit diese Anforderung sichergestellt ist, können verschiedene Testungen durchgeführt werden.

Es bestehen viele Verfahren zur Durchführung einer Usability-Testung. Eine Studie zur Evaluierung der entwickelten mobilen Applikation wurde durchgeführt und hier vorgestellt. Es werden Aspekte zur Planung, Vorbereitung und Durchführung einer Benutzerstudie beschrieben. Eine Auswahl der verschiedenen Methoden der Usability-Testung wird vorgestellt und verschiedene Szenarien der Durchführung diskutiert. Ziel der Studie ist es, die Benutzbarkeit der mobilen Applikation zu testen und die verschiedenen Darstellungen der Antwortmöglichkeiten in Fragebögen zu vergleichen.

Inhaltsverzeichnis

1 Einführung	1
1.1 Aufbau der Arbeit	3
2 Grundlagen	4
2.1. Mobile Datenerhebung	4
2.2. Mobile Applikation	5
2.3. Usability	6
2.3.1 Usability mobiler Applikationen	10
2.4. Questionnaire.....	10
3 Studien Design	12
3.1. Planung der Studie.....	12
3.2. Hypothesen.....	14
3.3. Vergleich der verschiedenen Usability-Evaluierungsmethoden	15
3.3.1 Datenerhebung.....	16
3.3.2 Standardisierte Usability-Fragebögen.....	17
3.3.3 Mögliche Szenarien der Benutzerstudie	19
4 Durchführung der Studien	20
4.1. Vorbereitung	20
4.1.1 Auswahl der Teilnehmer	20
4.1.2 Gestaltung der Aufgaben.....	24
4.2. Ablauf und Durchführung der Studie	25
4.3. Musterfragebogen	25
4.3.1 Demographie Fragebogen.....	26
4.3.2 System Usability Scale (SUS)	26
4.4. Lautes Denken.....	26

4.5. Vergleich der Design-Elemente in Fragebögen	28
5 Evaluierung.....	29
5.1. Demographische Daten.....	29
5.2. Ergebnisse	30
5.2.1 Lautes Denken-Methode	30
5.2.2 Vergleichsfragebogen.....	35
5.2.3 Auswertung des SUS-Fragebogen	39
5.3. Auswertung	41
6 Zusammenfassung und Ausblick.....	43
6.1. Ausblick	44

1

Einführung

„It is far better to adapt the technology to the user than to force the user to adapt to the technology.“ - Larry Marine

Software hat eine zunehmende Bedeutung in unserem Leben. Sie stellt einen integralen Bestandteil unseres Alltags dar, wird in immer neuen Anwendungsgebieten eingesetzt und unterstützt immer schwierigere Aufgaben [1]. Dabei spielt die Qualität der Software eine große Rolle. Da Software in den Gebieten wie Luftfahrt oder Automobilindustrie zum Einsatz kommt, müssen bei der Entwicklung von Software verschiedene Merkmale zur Qualitätssicherung beachtet werden. Es existieren mittlerweile einige Qualitätsmodelle, die diese Merkmale zusammenfassen. Eines davon ist das Qualitätsmodell nach ISO 25010, welches das frühere Modell nach ISO 9126 abgelöst hat [2] (Abb.1.1).

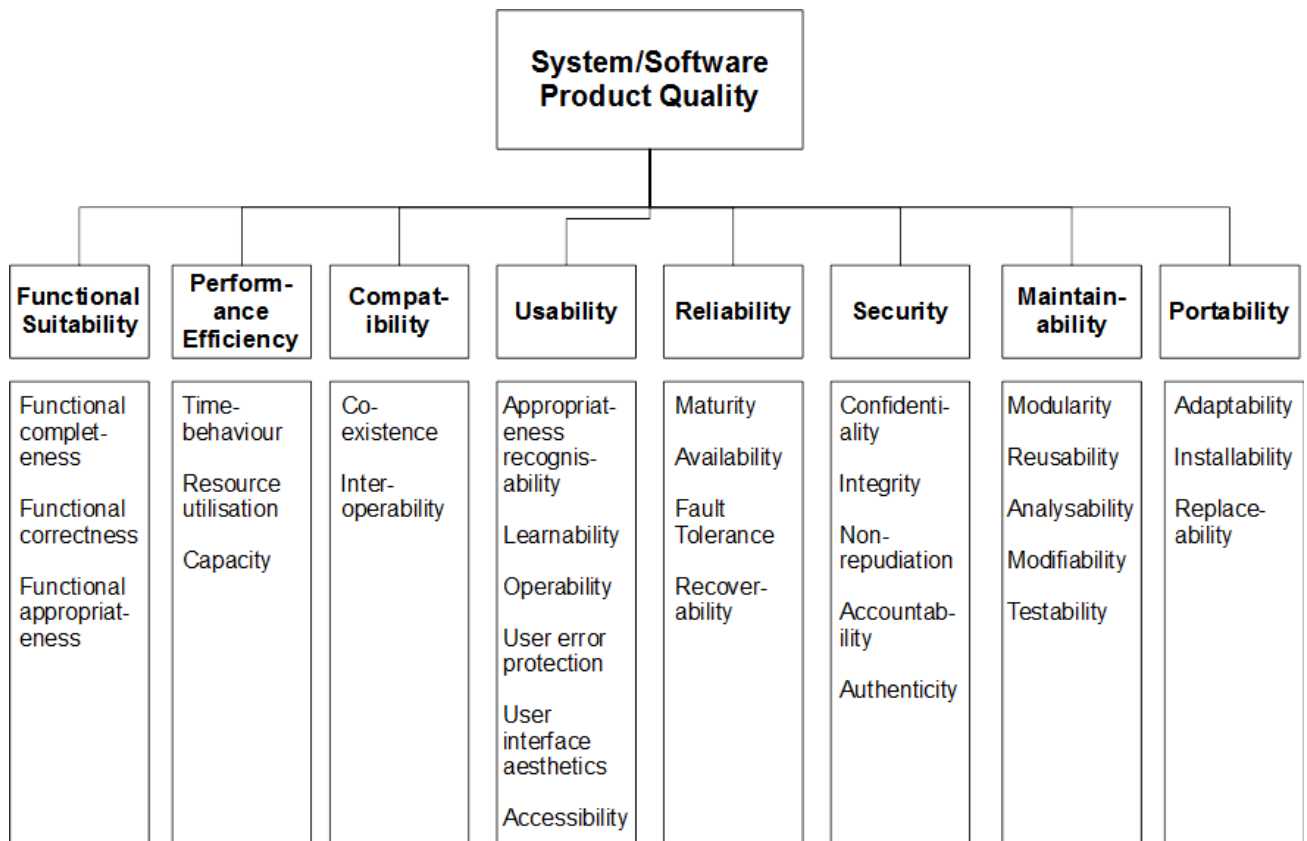


Abbildung 1.1: Qualitätsmodell für externe und interne Qualität nach ISO 25010 [3]

Das Modell der ISO 25010 zur Definition der externen und internen Qualität eines Softwareproduktes ist in acht Merkmale unterteilt, die sich weiter in Teilmerkmale aufteilen lassen. Ein Merkmal des Modells ist *Usability*.

Unter Usability versteht man ein Maß für Effektivität, Effizienz und Zufriedenheit mit dem Benutzer mit einem System vorgegebene Ziele erreichen können [4]. Diese Merkmale lassen sich durch Vorgehens- und Prozessmodelle, wie das V-Modell [5] nicht erreichen. Hier ist es notwendig, Usability-Testungen durchzuführen, um diese Merkmale festzustellen und korrigierend einzugreifen.

Die Aufgabe dieser Bachelorarbeit besteht darin, die im Rahmen des QuestionSys Projekts entwickelten mobilen Anwendung auf Usability zu testen. Dazu gehört die Bewertung des aktuellen Layouts, Darstellung der Fragebögen, sowie der Vergleich verschiedener Design-Elemente zur Darstellung der Antwortmöglichkeiten. Die bereits existierende Anwendung soll zur Darstellung und Bearbeitung von Fragebögen auf verschiedenen mobilen Endgeräten dienen. Außerdem sollen verschiedene Darstellungen der Antwortmöglichkeiten bei der Erstellung der Fragebögen benutzt und diese später

bearbeitet werden. Zum Zeitpunkt dieser Arbeit konnten noch keine veröffentlichte Evaluierung zur Usability der mobilen Anwendung ermittelt werden. Daher beabsichtigt diese Arbeit die Durchführung einer Studie mit dem Ziel die Benutzbarkeit zu testen.

Während der Entwicklung eines Softwareprodukts ist es wichtig auf diese Merkmale zu achten, damit das fertige Produkt von den Nutzern angenommen und benutzt wird. Durch Usability-Testungen ist es dem Entwickler möglich, das Produkt zu verbessern und es den Bedürfnissen der Nutzer anzupassen.

1.1 Aufbau der Arbeit

Im zweiten Kapitel werden die Begriffe *mobile Datenerhebung* und *mobile Applikation* erläutert. Anschließend wird die Bedeutung und Relevanz von *Benutzbarkeit* bei der Softwareentwicklung beschrieben. Zusätzlich wird der Unterschied von Usability im Zusammenhang mit mobilen Applikationen erläutert. Die in der durchgeführten Benutzerstudie evaluierte Applikation wird hier ebenfalls vorgestellt. Das dritte Kapitel behandelt die wichtigen Aspekte zur Planung einer Benutzerstudie. Dabei wird der Rahmen der Studie und die aufgestellten Hypothesen vorgestellt und die verschiedenen Methoden zur Usability-Evaluierung beschrieben und verglichen. Das vierte Kapitel beschäftigt sich mit der eigentlichen Vorbereitung und Durchführung der Benutzerstudie. Es wird auf die Faktoren eingegangen, die bei der Vorbereitung der Studie relevant sind. Der Ablauf und die Durchführung der Studie wird Schritt für Schritt diskutiert und die verwendeten Materialien und Instrumente werden beschrieben. Im Kapitel fünf werden die Ergebnisse der Benutzerstudie diskutiert. Die aus der Benutzerstudie gewonnenen Ergebnisse werden noch einmal zusammengefasst und ein Ausblick auf die möglichen Änderungen der Applikation gegeben.

2

Grundlagen

In diesem Kapitel werden die grundlegenden Begriffe wie *Datenerhebung* und *Usability*, auf die sich diese Arbeit bezieht, genauer erklärt. Zudem wird auf Begriffe wie *mobile Datenerhebung* und *Usability mobiler Applikationen* eingegangen. Die Applikation, die in der Benutzerstudie getestet wurde, wird hier auch vorgestellt.

2.1. Mobile Datenerhebung

Data Collection (auf Deutsch „Datenerhebung“) ist ein Prozess, in dem Informationen aus verschiedenen Quellen gesammelt und ausgewertet werden, um das Wissen oder Interesse in einem bestimmten Thema zu vervollständigen oder zu bestätigen [6].

Grundsätzlich bestehen verschiedene Methoden zur Datensammlung. Die weitverbreitetsten Methoden sind Fragebögen oder Interviews. Fragebögen werden üblicherweise auf Papier dargestellt, obwohl die digitale Variante immer mehr verbessert wird. Heutzutage versucht man sich vollständig von der Papiervariante zu lösen, da die digitale Variante deutlich effizienter und kostengünstiger ist [7]. So werden auch nach und nach Fragebögen in digitaler Form umgesetzt. Bei dem Versuch Zeit einzusparen, werden die Fragebögen dem Benutzer so simpel wie möglich zur Verfügung gestellt, sodass ein Fragebogen schnell ausgefüllt werden kann. Meistens versucht man hierbei die Benutzer durch Belohnungen zur Teilnahme zu motivieren und ihnen Instruktionen bezüglich der Bearbeitung des Fragebogens dann per E-Mail zukommen zu lassen. Üblich sind gebräuchliche Umfragen zur Kundenzufriedenheit, die einen wichtigen Beitrag in der Marktforschung leisten. Die andere Methode der Datensammlung ist die Durchführung von Interviews. Die meisten

Interviews werden telefonisch oder direkt mit dem Teilnehmer vor Ort durchgeführt. Interviews haben gegenüber Fragebögen den Vorteil, dass man einen erhöhten Prozentsatz an beantworteten Fragen feststellen kann. Bei Fragebögen kann es passieren, dass die Teilnehmer zu neutralen Antwortmöglichkeiten tendieren oder die Frage erst gar nicht beantworten. Hier handelt es sich um *Tendenz der Mitte* - ein Konstrukt der Fragebogenforschung, welches vorliegt, wenn Extremurteile weitgehend vermieden und in erster Linie mittlere Ausprägungsgrade ausgewählt werden [8]. Dies lässt sich durch direkte Kommunikation mit dem Interviewten vermeiden. Jedoch muss man hier ebenfalls bedenken, dass man für die Interviews deutlich mehr Zeit und somit auch Kosten einplanen muss.

Zur Aufnahme eines Interviews oder digitaler Darstellung von Fragebögen werden auch mobile Endgeräte genutzt. Ein mobiles Endgerät stellt ein elektronisches, tragbares Endgerät dar, das für Sprach- und Datenkommunikation genutzt wird.

Mobile Datensammlung ist eine Methode der Datensammlung über ein mobiles Endgerät, wie zum Beispiel ein Smartphone oder ein Tablet. Mobile Datensammlung kann überall genutzt werden, wo Daten gesammelt werden sollen. Eine Möglichkeit auf mobilen Endgeräten Daten zu sammeln, ist über Applikationen [9]. Dies soll nun im Folgenden genauer erläutert werden.

2.2. Mobile Applikation

Eine mobile Applikation, besser bekannt unter der Abkürzung *App*, ist eine Software, die auf einem mobilen Endgerät installiert werden kann. Applikationen stellen kleinere Softwareeinheiten mit begrenzten Funktionen dar. Diese Funktionen beziehen sich meistens auf Erreichbarkeit, Zugänglichkeit und Standorterkennung. Applikationen werden mittlerweile für jedes Anwendungsgebiet entwickelt. Je nach Anforderungen, die für die Entwicklung der Applikationen gestellt werden, lassen sich diese in native und plattformunabhängige Applikationen aufteilen.

Native Applikationen werden gezielt für eine bestimmte Plattform entwickelt. Bei mobilen Endgeräten ist diese Plattform vom Hersteller abhängig. Die Applikation wird nur für das jeweilige Betriebssystem des mobilen Endgeräts entwickelt und kann dadurch optimal auf die Hardware-Funktionen zugreifen.

Plattformunabhängige Applikationen laufen auf verschiedenen Plattformen. Sie lassen sich weiter in Web- und Hybrid-Applikationen unterteilen.

Web-Applikationen sind nicht auf dem mobilen Endgerät installiert, sondern lassen sich über den Webbrowser des mobilen Endgeräts aufrufen.

Hybride-Applikationen vereinen die Vorteile von nativen und Web-Applikationen. Die native Komponente kann direkt auf die betriebssystemspezifischen Funktionalitäten zugreifen, während die Web-Komponente in einem entsprechenden Container läuft [10].

2.3 Usability

Softwareentwicklung ist aus der heutigen Welt nicht mehr wegzudenken. Die Ergebnisse sind jeden Tag auf dem Markt zu beobachten. Applikationen werden ständig weiterentwickelt. Viele Faktoren spielen dabei eine wichtige Rolle. Einige grundlegende Anforderungen bei der Softwareentwicklung sind [7]:

- Verstehen des Softwareproduktes/Prozesses
- Definition der Qualität des Produktes/Prozesses
- Evaluierung der Erfolge und Misserfolge
- Informationsfeedback über den Projektverlauf
- Lernen aus Erfahrung
- Wiederverwendung von relevanten Erfahrungen

Eine wichtige Rolle spielen dabei Experimente und empirische Studien. Der Hauptgrund für die Durchführung von Experimenten und empirischen Studien in der Softwareentwicklung ist das Erzielen von objektiven und statistisch signifikanten Ergebnissen bezüglich Verständnis, Kontrolle, Vorhersagen und Weiterentwicklung des Produktes, bzw. des Prozesses.

Abhängig vom Zweck der Evaluierung und den Voraussetzungen der Untersuchung, kann zwischen drei großen verschiedenen, empirischen Methoden unterscheiden werden: der Umfrage, der Fallstudie und dem Experiment. Inwiefern ein Softwareprodukt den oben genannten Anforderungen entspricht, kann durch diese Methoden festgestellt werden.

Eine Umfrage ist eine Methode, um Informationen über Wissen, Standpunkte und Verhaltensweisen von oder über Menschen zu erhalten. Sie kann über Interviews oder Fragebögen erfolgen. Die Umfrage stellt, im Bereich der Marktforschung, ein wichtiges Instrument zur Datenerhebung dar [7].

Der Begriff der Fallstudie stellt in der Softwareentwicklung eine empirische Untersuchung dar, die zeitgenössische Phänomene der Softwareentwicklung innerhalb ihres realen Kontext untersucht, insbesondere wenn die Grenzen zwischen Phänomen und Kontext nicht eindeutig sind [11].

Ein Experiment (oder kontrolliertes Experiment) stellt eine empirische Methode da, in der eine Variable verändert wird und alle anderen Variablen konstant gehalten werden, um kausale Zusammenhänge zu isolieren. Experimente werden meistens im Labor unter kontrollierten Bedingungen, durchgeführt.

Weitere empirische Methoden zum Testen der Anforderungen, wie Effektivität, Effizienz und Zufriedenheit des Softwareproduktes, sind Usability-Evaluationsmethoden. Es sprechen viele Aspekte dafür, das Produkt einer Usability-Testung zu unterziehen. Usability-Tests können dabei helfen Design-Fehler zu entdecken oder die Verständlichkeit der Darstellung zu verbessern. Durch Usability-Tests kann der Entwickler bzw. Forscher sehen, wo die Benutzer Probleme beim Gebrauch des Produktes oder der Software haben und kann diese somit beheben. Ein anderer Vorteil der Usability-Tests ist, dass sie ohne hohe Kosten oder großen Zeitaufwand durchführbar sind. Für jeden Entwickler, Forscher oder Hersteller ist es ratsam eine Usability-Testung durchzuführen. Der Benutzer wird das Produkt in jedem Fall testen und wenn er – bereits nach Abschluss der Entwicklungsphase - Fehler entdeckt, werden diese deutlich mehr Zeit und Geld kosten, als wenn man sie bereits in der frühen Phase der Entwicklung und Fertigstellung entdeckt hätte. Um zu verstehen wieso es wichtig ist solche Testungen durchzuführen, wird hier genauer auf den Begriff *Usability* eingegangen.

Der englische Begriff „Usability“ wird häufig mit „Benutzbarkeit“, „Gebrauchstauglichkeit“ und „Benutzerfreundlichkeit“ übersetzt. Für Usability gibt es viele Definitionen, wie dieser Begriff genau beschrieben werden soll und was er alles umfasst.

- Die Definitionen laut Herceg (2005) besagt [12]: „Usability, auf Deutsch Gebrauchstauglichkeit, beschäftigt sich als Teilgebiet des Software Engineerings mit Theorien und Methoden zur Konzeption, Realisierung und Testung von benutzergerechten Anwendungssystemen“.

- „Usability is a quality attribute relating to how easy something is to use. More specifically, it refers to how quickly people can learn to use something, how efficient they are while using it, how memorable it is, how error prone it is and how much users like using it.” [13]
- “Usability bezeichnet das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und mit Zufriedenheit zu erreichen.” [14]

Der Begriff Usability beschränkt sich nicht auf eine Bedeutung, sondern wird auch mit folgenden anderen Begriffen assoziiert und zeichnet sich dadurch durch 5 folgende Merkmale aus [15]:

1. Lernfähigkeit: Wie einfach ist es für den Benutzer die grundlegenden Aufgaben zu erledigen bei der ersten Begegnung mit dem System?
2. Effizienz: Wenn der Benutzer einmal die Struktur und Bedienung kennt, wie effizient kann er mit dem System umgehen?
3. Einprägsamkeit: Wenn der Benutzer nach einer Zeit das System wiederverwendet, wie einfach kann er den Umgang damit wiederherstellen?
4. Fehler: Wie viele Fehler macht der Benutzer, wie gravierend sind diese Fehler und wie schnell kann er den Fehler beheben?
5. Zufriedenheit: Wie zufrieden ist der Benutzer mit der Benutzung des Systems?

Es gibt verschiedene Arten, Usability zu testen. Jedoch gibt es Methoden, die kostengünstiger sind wie beispielsweise die Beobachtung, Szenarien, „Lautes Denken“-Protokoll, heuristische Evaluation.

Bei der Beobachtung der Benutzer handelt es sich um ein einfaches Prinzip der Forschung, das man auf verschiedene Arten durchführen kann. Ein Beispiel wäre die Beobachtung von Personen bei realen Gegebenheiten. Die wichtigste Regel dabei ist, die Benutzer zu beobachten, sich still zu verhalten und die Benutzer sich ganz normal verhalten lassen, wie sie es üblicherweise machen würden.

Szenarien sind eine einfache Möglichkeit, einen bestimmten Teil des Produktes zu testen. Man kann auf den verschiedenen Ebenen der Entwicklung den Prototypen testen lassen. Wenn das Szenario klein ist, kann man es schnell wechseln oder verändern. Somit erreicht man, dass die Benutzer in kurzer Zeit öfters ein Feedback abgeben.

Die „Lautes Denken“ Methode wird in einem späteren Kapitel genauer beschrieben. Es handelt sich dabei um Einzelsitzungen, in denen der Benutzer eine Aufgabenstellung erhält, die er mit dem System durchführen und dabei seine Gedanken dazu laut äußern soll. Diese Methode hilft dem Versuchsleiter, bzw. Forscher nicht nur herauszufinden was der Benutzer mit dem System macht, sondern auch warum. Dieser Einblick in die Gedanken des Benutzers hilft festzustellen, welche Probleme das System aufweist und wo es genau zu Missverständnissen kommt.

Für die heuristische Evaluierung wird ein kleines Team von Experten als Gutachter zusammengestellt. Die Experten untersuchen das Produkt anhand von Usability-Richtlinien, um Probleme des Produktes zu finden. Jeder Gutachter untersucht das Produkt alleine. Nachdem alle Untersuchungen abgeschlossen sind, teilen sich die Experten ihre Ergebnisse mit. Das Resultat ist eine Liste mit allen gefundenen Usability-Problemen, die eine der Richtlinien verletzen. Die Probleme sind dann im Bezug auf die Richtlinien genau beschrieben.

Um Usability zu untersuchen ist die wahrscheinlich hilfreichste Methode die Durchführung einer Benutzerstudie. Sinn einer Benutzerstudie ist es, eine gewisse Sache oder ein Phänomen in einer bestimmten Zeitspanne zu testen. Üblicherweise bedienen dabei tatsächliche Nutzer das zu evaluierende Produkt oder System und werden in der Regel dabei vom Usability-Experten beobachtet. Dazu wird dem Nutzer eine realitätsnahe Situation und Aufgabe dargestellt, die er mit Hilfe des Produktes, bzw. des Systems lösen soll. Der Experte versucht dann, während der Studie, detaillierte Informationen zu den bestimmten Fragestellungen zu sammeln. Eine Benutzerstudie muss ganz genau geplant werden. Wie man eine solche Studie plant, folgt in den weiteren Kapiteln.

2.3.1 Usability mobiler Applikationen

Usability mobiler Applikationen unterscheidet sich von der Usability anhand der Computersysteme. Eine Rolle spielt dabei die unterschiedliche Architektur der zwei Systeme, wobei auf die Bildschirmgröße, Dateneingabe, Kapazität und Leistung geachtet werden sollte. Die weit verbreiteten Apps, wie die von Google oder Apple, sind beim Kauf des Smartphones auch gleich verfügbar auf dem Gerät.

Die Benutzer tendieren dazu, die Applikationen herunterzuladen, die leicht zu bedienen sind, bzw. wenig Zeit kosten, um eine Aufgabe zu erledigen und zugleich

benutzerfreundlich bedienbar sind. Auch bei mobile Applikation Usability benutzt man die gleichen Methoden der Evaluation wie bei Usability-Testung von Produkten/Systemen. In dieser Arbeit wird anhand einer Benutzerstudie erklärt und beschrieben, wie man Usability auf einer mobilen Anwendung testet. Für die Benutzerstudie wurde die Applikation *Questionnaire* getestet.

2.4. Questionnaire

Die Applikation *Questionnaire* (auf Deutsch: *Fragebogen*) ist nur ein Teil eines Fragebogensystems namens *QuestionSys*. *QuestionSys* wurde entwickelt, um die digitale Darstellung der Fragebögen an mobilen Endgeräten und die Datensammlung zu erleichtern und generell jede Phase der Datensammlung zu unterstützen. Es ist eine große Arbeit die Fragebögen von Papier wieder zu digitalisieren, um die Daten zu speichern. Natürlich kann es bei der Übertragung auch zu Fehlern bei der Eingabe kommen. Solche Fehler wären fatal, wenn man sich vorstellt, dass es z.B. um Patientenakten geht. Die Anwendung soll dabei helfen, die Fragebögen ganz einfach zu erstellen und danach in ein ausführbares Prozessmodell umzuwandeln. Diese Prozessmodelle können dann auf das mobile Endgerät übertragen und z.B. auf einem Smartphone benutzt werden. Die gesammelten Daten werden an einer Stelle gespeichert und erleichtern so die Datenerhebung.

Das *QuestionSys* Framework sieht folgendermaßen aus [16]:

1. Modellierung des Instruments zur Datensammlung (Konfigurator)
2. Mapping
3. Prozessmodell
4. Ausführen auf mobilen Endgeräten (*Questionnaire* App)
5. Middleware für sichere Datenübertragung und Datenverwaltung

Bei der Modellierung wird das Instrument zur Datensammlung (Fragebogen) erstellt und zusammengebaut. In der Anwendung gibt es verschiedene Auswahlmöglichkeiten wie die Fragen und Antwortmöglichkeiten aussehen sollen. Es werden Seiten erstellt, auf denen man dann Titel, Beschreibung des Fragebogens und die Fragen selbst einfügen kann. Für die Antwortmöglichkeiten kann man zwischen verschiedenen Darstellungen auswählen, wie beispielsweise zum Ankreuzen (*Single-Choice* / *Multiple-Choice*), als freies Textfeld oder als Slider. Ist man zufrieden mit dem fertigen Fragebogen, wird dieser in ein Prozessmodell

umgewandelt und kann dann auf das Endgerät exportiert werden. Nach dem Einsatz der Fragebögen und dem erfolgreichen Sammeln der Daten, werden diese sicher übertragen und gespeichert.

Die Entwickler haben bei dem Projekt darauf geachtet, dass folgende Voraussetzung für die mobile Anwendung erfüllt sind [16]:

- Offline Ausführbarkeit zur Verfügung stellen
- Prozess Flexibilität ermöglichen: z.B. die Reihenfolge der Fragen muss der Verständlichkeit wegen, flexibel zur ändern sein
- Anpassbare Benutzerschnittstelle zur Verfügung stellen

Durch diese Eigenschaften ist das System weitreichend einsetzbar. Änderungen können leicht durchgeführt werden und die Benutzerschnittstelle ist den Interessen des Benutzers anpassbar. Ein weiterer Vorteil der Anwendung ist, dass die gesammelten Daten verschiedener Versionen der erstellten Fragebögen nicht manuell zusammengeführt werden müssen.

Bei der Entwicklung kann jedoch nicht alles bedacht und jeder Fehler entdeckt werden. Deswegen ist es hier, wie auch bei jedem Produkt oder Programm notwendig, eine Testung durchzuführen. Bei der hier vorgestellten Studie wurde die Testung der frühen Phasen der Entwicklung erfolgreich abgeschlossen. Hier wird ausschließlich die Benutzbarkeit der App mit schon fertiggestellten Fragebögen getestet. Die größte Aufmerksamkeit gilt der Usability. Zusätzlich werden die verschiedene Aspekte der Anwendung beachtet, um aufzudecken ob und wenn ja, welche optimiert werden sollten.

3

Studien Design

In diesem Kapitel werden Möglichkeiten zur Durchführung einer Benutzerstudie vorgestellt. Zudem wird erläutert, welche Faktoren in die Planung einfließen. Außerdem werden der Plan der durchgeführten Benutzerstudie und die aufgestellten Hypothesen vorgestellt.

3.1 Planung der Studie

Die Durchführung einer Studie besteht aus 5 Schritten [7]:

1. Design der Studien: Das Objekt welches getestet wird, wird definiert und die Studie wird geplant
2. Vorbereitung der Datensammlung: Prozeduren und Protokolle für die Datensammlung werden erstellt
3. Datensammlung: Sammeln der Daten während der Studie
4. Auswertung: Die gesammelten Daten werden ausgewertet
5. Ergebnisse: Ergebnisse der Studie werden zusammengefasst und

Dieser Prozess ist bei empirischen Studien fast immer identisch.

Eine gute Planung ist ausschlaggebend für eine erfolgreiche Studie. Bevor man die Studie planen kann, ist es wichtig das Ziel der Studie zu definieren. Das geschieht mit der Bestimmung des Umfangs der Studie. Um das zu erleichtern, wurden in [7] die GoalQuestionMetric Methode entwickelt (kurz GQM). Die GQM-Methode baut auf der Annahme auf, dass eine Organisation oder ein Unternehmen folgende Aspekte erfüllen

muss, um zweckmäßig richtige Messungen durchzuführen: *Ziele* für sich selbst und für das Projekt definieren; die bestehenden *Daten* nutzen, um das Projekt zu charakterisieren und einen *Rahmen* des Projektes und der Daten zur Verfügung stellen, mit Berücksichtigung der definierten Ziele. Mit diesen Anforderungen soll als Ergebnis ein Messmodell zur Verfügung stehen, welches Probleme herausfiltern und Regeln zur Messung definieren soll. Das Messmodell besteht dabei aus 3 Ebenen [7]:

1. Goal (Begriffsebene): Ein Ziel für ein bestimmtes Objekt, mit Berücksichtigung verschiedener Gründe, Qualitätsmerkmale, Perspektiven und der Umgebung wird definiert.
2. Question (Operationsebene): Fragen werden festgelegt, um das Objekt, welches getestet wird, zu charakterisieren und um herauszufinden, wie die definierten Ziele erfolgreich erfüllt werden können.
3. Metric (Quantitative Ebene): Daten, die in Verbindung mit den Fragen stehen, werden so festgelegt, dass sie auf eine quantitative Weise die Fragen beantworten, heißt weder objektiv noch subjektiv.

Der aufwändigste Teil ist es die Ziele sinnvoll zu definieren. Ist das erledigt, können dann die passenden Fragen abgeleitet und anhand dieser schließlich die Metriken definiert werden.

Nachdem der Rahmen mit dem Ziel, was gemacht und wie es gemessen werden soll festgelegt wurde, kann zum nächsten Schritt übergegangen werden: der Planung. Der Rahmen sagt nur aus, wieso die Studie durchgeführt werden soll, bei der Planung wird schließlich festgelegt, wie sie konkret durchgeführt werden soll. Anhand der Zielsetzung kann der Umfang definiert werden, wie zum Beispiel die Umgebung, in der die Studie durchgeführt werden soll. Als nächstes werden die Hypothesen formuliert, die durch die Studie bestätigt oder widerlegt werden sollen. Für die Messung werden die passenden Instrumente ausgewählt und zum Schluss hilft die Evaluierung der Validität, ob die Studie wirklich das gemessen hat, was sie sollte.

Die im Rahmen dieser Arbeit durchgeführte Studie ist eine Benutzerstudie. In der Studie sollte eine Applikation auf allgemeine Benutzbarkeit getestet werden. Zusätzlich wurden die grafischen Darstellungen der Antwortmöglichkeit mit den Antwortmöglichkeiten zum Ankreuzen verglichen. Ziel der Studie war es, die Gebrauchstauglichkeit und das derzeitige

Interface der Applikation *Questionnaire* zu testen und die erhobenen Daten und mögliche Verbesserungsvorschläge zusammenzufassen. Im Gespräch mit dem Entwickler der Applikation, wurde ein Plan für die Testung zusammengefasst. Bei der Planung der Studie wurden Aspekte wie Umgebung, Durchlauf, Instrumente und Metriken zur Bewertung festgelegt. Die Studie sollte einen Durchlauf mit verschiedenen Bearbeitungsschritten beinhalten, der zwischen 20 und 30 Minuten pro Teilnehmer beanspruchen sollte. Bei der Anzahl der Teilnehmer wurde die Zahl auf mindestens 25 Teilnehmer gesetzt, wobei die Voraussetzung zur Teilnahme an der Studie Kenntnisse im Umgang mit mobilen Endgeräten war. Für das Testen der Applikation wurde ein Tablet mit dem Android-Betriebssystem zur Verfügung gestellt, worauf die Applikation *Questionnaire* installiert war. Zusätzlich wurde die Applikation *AZ Screen Recorder* installiert, die für die Aufnahme der Oberfläche und Audio benutzt wurde. Die Aufnahmen wurden für die Datenerhebung benötigt. In der Applikation *Questionnaire* wurden durch die verschiedenen Visualisierungen der Fragebögen die neuen Design-Elemente angezeigt und getestet. Der Teilnehmer kann dadurch die Applikation aktiv testen und seine Eindrücke äußern.

Während der Benutzerstudie sollten Daten gesammelt werden [17] wie: erfolgreich abgeschlossene Aufgaben, Fehler, die durch den Teilnehmer oder die Applikation verursacht wurden, Zeit, die der Teilnehmer benötigt hat, um die Aufgabe abzuschließen, Zufriedenheit der Teilnehmer zur Bedienbarkeit, Darstellungen und Hilfestellungen der Applikation und positive, negative Kommentare und Verbesserungsvorschläge.

3.2 Hypothesen

Von Interesse der Studie sind, neben der allgemeinen Benutzbarkeit der Applikation, auch der Eindruck der neuen Design-Elemente zur Darstellung der Antwortmöglichkeiten. Dazu wurden Aussagen festgelegt die durch die Studie bestätigt oder widerlegt werden sollten.

Folgende Hypothesen wurden formuliert:

H0a: Die neuen Design-Elemente als Antwortmöglichkeiten, sind gleich gut wie die bekannten Antwortmöglichkeiten zum Ankreuzen.

H0b: Die neuen Design-Elemente werden von den Teilnehmern eher bevorzugt als die bekannten Antwortmöglichkeiten zum Ankreuzen.

H1a: Die Teilnehmer empfinden es genauso einfach, den Fragebogen mit neuen Design-Elementen auszufüllen, wie den mit den bekannten Antwortmöglichkeiten zum Ankreuzen.

H1b: Die Teilnehmer empfinden es einfacher, den Fragebogen mit Antwortmöglichkeiten zum Ankreuzen auszufüllen, als den mit verschiedenen Design-Elementen wie Slider, freies Textfeld oder Range Elementen.

H2a: Die Teilnehmer finden das Ausfüllen der Fragebögen auf dem mobilen Endgerät gleich gut wie auf Papier.

H2b: Die Teilnehmer bevorzugen das Ausfüllen der Fragebögen auf dem mobilen Endgerät, als auf Papier.

3.3 Vergleich der verschiedenen Usability-Evaluierungsmethoden

Wenn der Benutzer mit einer Applikation arbeitet, möchte er diese einfach bedienen können. Bei der Entwicklung der Applikation ist, wie auch bei jedem anderem Produkt, darauf zu achten, dass die Bedienung nicht unnötig kompliziert ist, sondern einfach und intuitiv verständlich ist. Die Benutzer möchten heute mit wenigen Klicks an ihr Ziel kommen. Dafür ist einfache und verständliche Bedienbarkeit ein Muss. Ob diese Usability-Anforderungen erfüllt sind, lässt sich durch verschiedene Methoden feststellen. Diese reichen von Fragebögen und Interviews bis zu Experimenten im Labor, Feldstudien, usw. Da es sich in dieser Abschlussarbeit um eine Benutzerstudie handelt, wird hier genauer auf Methoden eingegangen, die dafür geeignet sind.

Die Auswahl der passenden Evaluierungsmethode ist von vielen Faktoren abhängig. Wichtig ist dabei darauf zu achten, dass die Evaluierungsmethode möglichst viele oder am besten alle Usability-Merkmale berücksichtigt. Dies kann allerdings sehr zeitaufwändig und komplex werden, wenn das Merkmal *Einprägsamkeit* [15] berücksichtigt werden soll. Dabei soll der Benutzer nach einiger Zeit den Umgang mit der Anwendung erneut durchführen.

Bei dem Usability-Merkmal *Zufriedenheit* [15] können Faktoren wie Umgebung, Wohlbefinden oder das mobile Endgerät eine Rolle spielen. Das Usability-Merkmal bezieht sich auf das Wohlbefinden und den Zufriedenheitsgrad des Benutzers während des Tests. Jedoch kann das Wohlbefinden, kurz vor der Evaluierung, auch in seine Bewertung miteinfließen. Fühlt sich der Benutzer zum Beispiel durch die Umgebung gestört oder hat

Probleme im Umgang mit dem Betriebssystem auf dem die Applikation installiert ist, kann sich dies negativ auf die Beurteilung auswirken.

Um die Voraussetzungen der Studie den realen Situationen anzupassen, in denen die Applikation später in Benutzung sein wird, wurde die Studie nicht in einem Labor durchgeführt. Die Teilnehmer der Studie könnten im Lernraum an der Universität oder Zuhause, in ihrer bekannten Umgebung an der Studie teilnehmen und die Applikation testen. Dadurch befanden sich die Teilnehmern in einer, für sie bekannten Umgebung, wodurch das Befinden der Teilnehmer während der Studie so angenehm wie möglich sein sollte.

3.3.1 Datenerhebung

Bei jeder Testung nimmt die Auswahl der Methode zur Datenerhebung großen Einfluss auf den Verlauf und Umfang der Testung ein. Laut [18], lassen sich die Methoden der Datenerhebung in drei Stufen unterteilen:

- Erste Stufe: Die direkte Methode bedeutet, dass der Forscher im direkten Kontakt zum Benutzer steht und Daten in Echtzeit erhebt. Das ist beispielsweise bei Interviews oder Beobachtungen mit der *Lauten Denken* Methode der Fall.
- Zweite Stufe: Bei den indirekten Methoden, sammelt der Forscher direkt Rohdaten, ohne dabei mit dem Benutzer während der Datenerhebung zu interagieren. Dazu zählen beispielsweise Beobachtungen mit Videoaufnahmen.
- Dritte Stufe: Unabhängige Analyse mit schon vorhandenen und verfügbaren Daten. Das wäre der Fall beim Analysieren von Daten von Lastenheften oder Fehlerberichten eines Unternehmens oder dem Auswerten von Daten wie z.B. Zeitmessungen aus der Datenbank des Unternehmens

Die erste Stufe der Datensammlung bedarf der meisten Kosten und Zeitaufwand, da die Ergebnisse vom Benutzer, genauso wie vom Forscher, abhängen. Vorteil der ersten und der zweiten Stufe der Datenerhebung stellt das große Ausmaß der Kontrolle dar. Der Forscher kann während der Testung bestimmen, welche Daten genau gesammelt werden sollen. Die dritte Stufe der Datenerhebung stellt die günstigste Methode dar, jedoch hat der Forscher wenig Einfluß auf die Qualität der Datenerhebung und muss mit den bereits

vorhanden Daten arbeiten, die möglicherweise nicht genau den Zweck der Studie erfüllen, sondern für einen anderen Zweck gesammelt wurden.

In der hier vorgestellten Benutzerstudie sollten die Daten mit Methoden der ersten Stufe erhoben werden. Da die Benutzerstudie einmalig, und nicht über längere Zeit durchgeführt wurde, standen die Daten sofort zur Auswertung bereit und wurden anschließend zusammengefasst.

3.3.2 Standardisierte Usability-Fragebögen

Für die Usability-Evaluierung wurden in den letzten Jahren verschiedene Fragebögen entwickelt. Standardisierte Fragebögen sollen zusätzlich helfen, subjektive Einstellungen und Erfahrungen von Benutzern zu messen. Da eine objektive Datenerhebung Werte zur Zeitmessung und Fehlerraten liefert, soll durch die Fragebögen die Zufriedenheit der Benutzer mit dem Produkt eingefangen werden. Einige der Fragebögen bieten dafür offene Antwortfelder an, über die es möglich ist Verbesserungsvorschläge zu äußern, oder ein Usability-Problem konkret zu beschreiben. Bei der Recherche der verschiedenen Fragebögen, werden drei davon sehr oft in der Literatur beschrieben: IsoNorm 9241/10, Isometrics und System Usability Scale (kurz: SUS).

Isonorm 9241/10 wurde 1993 veröffentlicht und besteht aus 35 Fragen, die nach den 7 Kriterien der allgemeinen, ergonomischen „Grundsätze zur Dialoggestaltung“ aus der DIN EN ISO 9241-10 (Europäisches Komitee für Normung 1995) unterteilt sind. Jedes Kriterium hat 5 Aussagen, die in einer Skala dargestellt sind. Die Antwortmöglichkeiten bestehen aus einer 7-stufigen Likertskala, die mit sehr negativ „(---)“ bis sehr positiv „(+++)“ bewertet werden kann [14].

Der Isometrics-Fragebogen ist auch nach 7 Kriterien aus dem DIN EN ISO 9241-10 unterteilt. Der Fragebogen besteht insgesamt aus 75 Aussagen und die einzelnen Kriterien sind in 7-12 Aussagen unterteilt. Die einzelnen Aussagen kann man auf einer 5-stufigen Likertskala beurteilen. Die erste Stufe der Ratingskala steht für „Stimmt nicht“ und die fünfte für „Stimmt sehr“. Es besteht auch die Möglichkeit, „keine Angabe“ anzukreuzen. Von dem Isometrics-Fragebogen existiert eine lange und eine kurze Version. Die lange Version wird für formative Evaluierung und die kurze für summative Evaluierung benutzt. Der

Unterschied zwischen der langen und der kurzen Version besteht darin, dass in der langen Version zu jeder Aussage eine Gewichtung gegeben werden kann und in der langen Version sollen die Teilnehmer ihre Antwort begründen und ihre Anmerkungen äußern [14].

Der System Usability Scale (SUS) Fragebogen wurde 1986 von John Brooke vorgestellt [19]. Er hat eine einfache 10-Fragen Usability-Skala erstellt, welche die subjektive Bewertung der Benutzbarkeit einer Anwendung einfängt. Der SUS Fragebogen besteht aus einer Likertskala, die aus fünf Feldern besteht. Das erste Feld steht für die Aussage „Stimme gar nicht zu“, das letzte Feld für die Aussage „Stimme voll zu“. Der Fragebogen besteht aus 10 Aussagen, wobei die erste Aussage eine positive ist und die nächste eine negative Aussage ist. Die Aussagen werden abwechselnd als positiv und negativ dargestellt. Somit soll verhindert werden, dass der Teilnehmer schnell, ohne viel über die Aussagen nachzudenken, den Fragebogen ausfüllt. Durch die Abwechslung muss der Teilnehmer jede Aussage genau durchlesen, um die Zustimmung oder Ablehnung anzukreuzen. Die vorgegebenen Aussagen decken viele Aspekte der Benutzbarkeit eines Systems ab, wie Hilfestellungen, Lernfähigkeit und Komplexität. Somit ist SUS ein zuverlässige und robuste Usability-Evaluationsmethode und kann in verschiedenen Bereichen der Industrie oder wissenschaftlichen Projekten eingesetzt werden [19].

Möchte man für die Usability-Evaluierung nur einen standardisierten Fragebogen verwenden, so sollte man auf Isonorm oder Isometrics zurückgreifen, da diese eine ausführliche Usability-Evaluierung nach Gütekriterien ergeben. Der SUS-Fragebogen ist deutlich kürzer im Vergleich zu den anderen Fragebogen, fasst aber die Merkmale von Usability gut zusammen. Laut [20] liefert der SUS-Fragebogen schon bei 8 bis 10 Benutzern verwendbare Ergebnisse.

3.3.3 Mögliche Szenarien der Benutzerstudie

Drei mögliche Szenarien, wie man die Benutzerstudie durchführen könnte, wurden entwickelt und werden hier vorgestellt.

Szenario 1: Eine Kontrollgruppe und eine Experimentgruppe sollen aufgestellt werden. Der Experimentgruppe sollen Fragebögen mit grafischer Darstellung der Antwortmöglichkeiten

zugeteilt werden und der Kontrollgruppe Fragebögen ohne grafische Darstellung der Antwortmöglichkeiten. Die Aufgabe soll darin bestehen, sich an einem mobilen Endgerät, durch zwei Fragebögen durchzuklicken, die identisch aufgebaut sind, aber mit verschiedenen Themen. Während die Teilnehmer sich durchklicken, sollen sie laut ihre Gedanken zu der Benutzung der Anwendung äußern („Lautes Denken“-Methode).

Szenario 2: Eine Experimentgruppe und eine Kontrollgruppe sollen aufgestellt werden. Die Experimentgruppe soll die Fragebögen mit grafischer Darstellung der Antwortmöglichkeiten bekommen und die Kontrollgruppe die Fragebögen ohne grafische Darstellung der Antwortmöglichkeiten. Die Aufgabe besteht aus zwei Fragebögen, durch die sich die Teilnehmer durchklicken sollen. Nach Ausfüllen des ersten Fragebogens am mobilen Endgeräten, bekommen die Teilnehmer einen Fragebogen zum Ausfüllen, wo Fragen zu der Darstellung der Fragen und Antwortmöglichkeiten gestellt werden. Der gleiche Vorgang passiert nach dem Ausfüllen des zweiten Fragebogens an dem mobilen Endgerät. Am Ende der Studie soll noch ein standardisierter Usability-Fragebogen ausgefüllt werden.

Szenario 3: Jeder Teilnehmer der Studie bekommt die Aufgabe, sich durch zwei identisch aufgebaute Fragebogen durchzuklicken. Die Fragebögen werden auf dem mobilen Endgerät ausgefüllt. Beide Fragebögen erhalten grafische Darstellungen der Antwortmöglichkeiten sowie Antwortmöglichkeiten zum Ankreuzen. Während sie die Fragebögen ausfüllen, sollen sie laut ihre Gedanken zu der Benutzbarkeit der Anwendung äußern („Lautes Denken“-Methode). Danach sollen die Teilnehmer einen Vergleichfragebogen ausfüllen, um die Darstellungen der Antwortmöglichkeiten zu bewerten.

Nach reiflicher Überlegung wurde das Szenario 3 ausgewählt. Diese Überlegung entsprach der effektivsten Methode zur Erfüllung der Ziele der Studie. Durch die Einzelsitzungen mit den Teilnehmern, die sich mit der Applikation beschäftigen, sollten mehr Daten über die Bedienung der Applikation gesammelt werden, als über Evaluierung durch Fragebögen. Durch die Vereinbarung der verschiedenen Darstellungen der Antwortmöglichkeiten, konnten die Teilnehmer alle möglichen Darstellungen in den zwei Fragebögen sehen und dadurch besser entscheiden, was sie bevorzugen.

4

Durchführung der Studie

Die Durchführung der Studie besteht aus 3 wesentlichen Schritten: der *Vorbereitung*, in der die Teilnehmer ausgewählt und Fragebögen vorbereitet werden; der *Durchführung*, in der die Teilnehmer an der Testung teilnehmen und die Daten erhoben werden und der *Datensvalidierung*, in der die erhobenen Daten geprüft werden, ob wirklich das gemessen wurde was beabsichtigt war [7].

In diesem Kapitel werden die Schritte der *Vorbereitung* und *Durchführung* genauer betrachtet. Außerdem wird die Struktur der Studie und die dafür verwendeten Materialien und Instrumente vorgestellt.

4.1 Vorbereitung

Bevor die eigentliche Studie durchgeführt wird, müssen einige wichtige Aspekte bei der Vorbereitung beachtet werden. Dazu zählen unter anderem die Auswahl der Teilnehmer und die Vorbereitung der Materialien, wie Fragebögen und Instrumente [7].

4.1.1 Auswahl der Teilnehmer

Bei der Auswahl der Teilnehmer sollte beachtet werden, dass diese auch der Zielgruppe entsprechen.

Abhängig vom Alter, stellen Nutzer unterschiedliche Anforderungen an das zu testende Produkt. Bedingung für die Teilnahme an der Studie waren Kenntnisse im Umgang mit mobilen Endgeräten, was bei der Zielgruppe (18-40 Jahre) auch kein Hindernis dargestellt hat.

Die Anforderungen, die an die Teilnehmer gestellt werden, stellen nur einen Aspekt bei Auswahl der richtigen Teilnehmer dar. Auch wenn eine Studie gut geplant ist und die richtigen Evaluierungsmethoden ausgewählt wurden, werden die Ergebnisse nicht valide sein, wenn der Teilnehmer die Studie nicht erst nimmt [7]. Trotz der klaren Aussage, dass nur das Produkt und nicht die Leistung des Teilnehmers getestet wird, kann sich bei dem Teilnehmer in so einer Testung das Unbehagen erhöhen, insbesondere wenn er glaubt etwas falsch gemacht zu haben.

Dadurch, dass die Auswahl der Umgebung, wo die Studie stattfinden soll, dem Teilnehmer überlassen wurde, wurde versucht die Situation für den Teilnehmer angenehmer zu gestalten. Ergänzend zu den Lehrräumen der Universität, kann es insbesondere durch den mobilen Kontext der Applikation sinnvoll sein, den Teilnehmer in seinem natürlichen Umfeld, zum Beispiel zuhause, an der Studie teilnehmen zu lassen. So kann festgestellt werden, ob die Applikation auch im mobilen Kontext problemlos nutzbar ist [22].

Der Versuchsleiter spielt beim Wohlbefinden der Teilnehmer während der Testung ebenfalls eine wichtige Rolle, indem er den Teilnehmer genügend Informationen zur Verfügung stellt, damit diese auch richtig an die Aufgabe herangehen. Er sollte auch beruhigend auf die Teilnehmer einwirken, damit sich diese, falls ein Problem auftritt, nicht verunsichert oder schuldig fühlen [15].

Die Ergebnisse sollten bei jeder Studie vertraulich behandelt werden. Dazu wurden in dieser Studie Codes vergeben, die eine nachträgliche Identifizierung der Teilnehmer ausschließen. Die Teilnehmer wurden ausdrücklich darauf hingewiesen, dass anhand der während der gesammelten Daten nicht auf ihre Identität zurückgeschlossen werden kann. Die Anzahl der Versuchsleiter sollte so gering wie möglich gehalten werden, da die Teilnehmer durch das Beobachten mehrerer Personen verunsichert werden könnten und infolgedessen mehr Fehler machen.

Der Versuchsleiter sollte sicherstellen, dass alle benötigte Instrumente für die Studie bereitstehen. Werden bei der Testung Fragebögen eingesetzt, die vom Teilnehmer selbst auszufüllen sind, sollten diese immer anonym sein, was im Gegenzug auch bedeutet, dass die Teilnehmer nicht kontaktiert werden können, falls Unklarheiten auftreten [7].

Während der Testung sollte sich der Versuchsleiter nicht einmischen und dem Teilnehmer selber die Lösung für eine Problemstellung finden lassen. Nicht nur, dass es zu valideren Ergebnissen führt, es verhindert auch, dass sich der Teilnehmer unwohl fühlt und denkt, dass er das Problem nicht selbständig lösen kann. Auf der anderen Seite, sollte der Versuchsleiter den Teilnehmer nicht ewig mit der Aufgabe ringen lassen [15]. In solchen Fällen, hat der Versuchsleiter durch Stellen von Fragen, dem Teilnehmer zu einer Lösung verholffen.

4.1.2 Gestaltung der Aufgaben

Die Aufgaben waren von Umfang her schnell zu bearbeiten, sodass sie das Zeitlimit des Durchlaufs nicht überschritten, aber auch nicht zu kurz, um trivial zu wirken. Die Aufgaben waren klar gestellt und hatten eine logische Reihenfolge. Ebenfalls wurde dem Teilnehmer die Aufgabenstellung in schriftlicher Form vorgelegt und stand ihm während der ganzen Testung zur Verfügung, sodass jeder Teilnehmer die gleiche Formulierung erhielt und er sich während der Testung darauf beziehen konnte [15]. Bei der Gestaltung der Aufgaben sollte beachtet werden, dass diese realitätsnah sind, der Testsituation entsprechen und die wesentlichen Funktionen der Produktes abdecken.

Für die Testung wurden zwei Fragebögen mit dem Konfigurator erstellt und auf das mobile Endgerät übertragen. Bei dem mobilen Endgerät handelt es sich um das Tablet *Nexus 7* vom Hersteller ASUS mit 7-Zoll-HD-Display und dem Betriebssystem Android 6.0.1. Auf das Tablet wurde zusätzlich die Applikation *AZ Screen Recorder* installiert. Sie zeichnet alles auf, was auf einem Android-Display dargestellt wird [24]. Durch die verschiedenen Einstellungen der Applikation, wie *Berührung anzeigen*, *Auflösung*, *Overlay-Informationen*, hat sich die Applikation als das richtige Tool für die Testung gezeigt.

Beide Instrumente, das Tablet und die zusätzlich installierte Applikation wurden hinzugezogen, um die Aufzeichnung der Interaktion des Teilnehmers und der Applikation aufzuzeichnen und somit die Testung der Applikation möglich zu machen.

4.2. Ablauf und Durchführung der Studie

Bei dieser Studie wurden Einzelsitzungen mit den Teilnehmern vereinbart. Hierbei wurden die Ergebnisse der Datenerhebung sofort erhalten und der Versuchsleiter war während des ganzen Ablaufs präsent, um auftauchende Fragen umgehend beantworten zu können [2].

Die Studie war folgendermaßen aufgebaut:

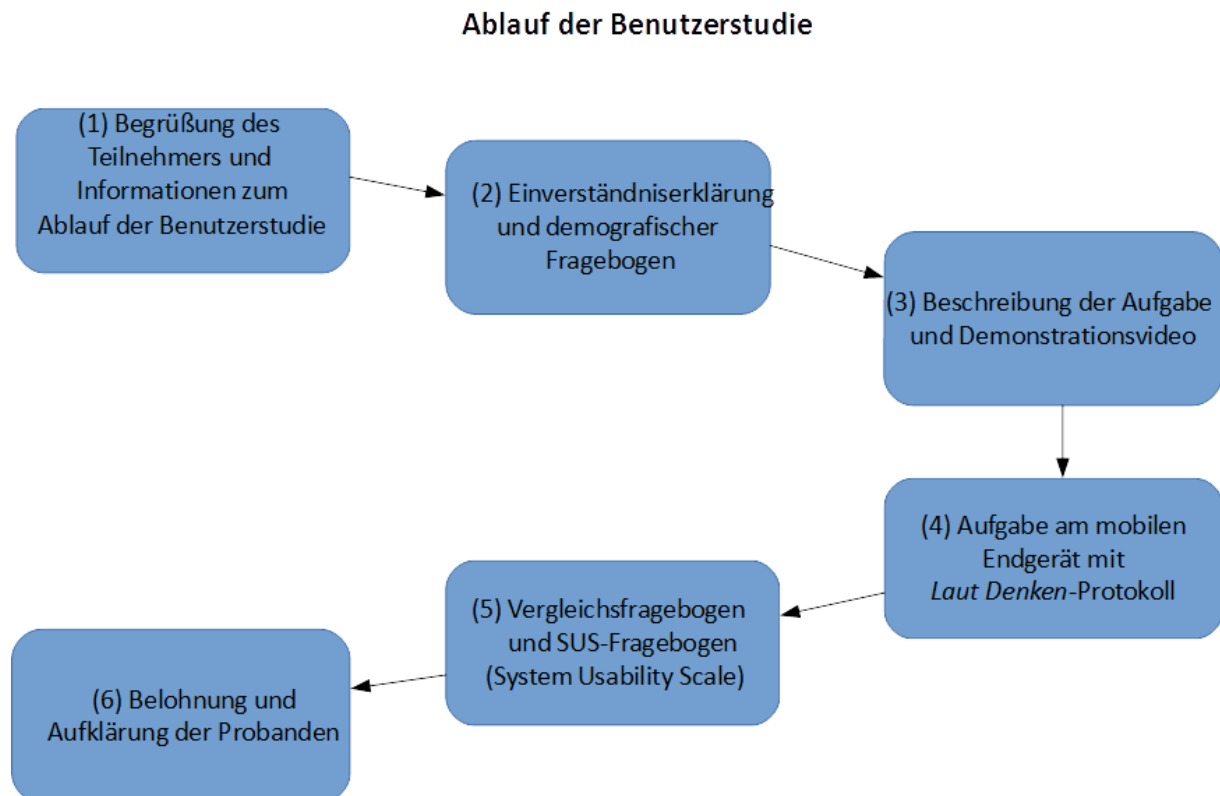


Abbildung 4.1: Ablauf der Benutzerstudie

Schritt 1 & 2: Informationen zum Ablauf, Einverständniserklärung und demografischer Fragebogen

Der Teilnehmer wurde vom Versuchsleiter begrüßt und in den Ablauf der Studie eingeweiht. Dem Teilnehmer wurde als erstes die Einverständniserklärung vorgelegt, die alle wichtigen Informationen zur Vertraulichkeit der gesammelten Daten und Ablauf der Studie erklärt. Nach dem Unterzeichnen der Einverständniserklärung, folgte ein demografischer

Fragebogen mit allgemeinen Fragen zur Person, wie Alter, Geschlecht, Bildungsgrad, Umgang mit mobilen Applikationen etc.

Schritt 3 & 4 : Instruktionen und Durchführung der Aufgabe am mobilen Endgerät

Danach folgte die Bearbeitung der Aufgabe am mobilen Endgerät, die durch das *Laut Denken*-Protokoll begleitet wurde. Der Teilnehmer wurde mit einer Beschreibung der Aufgabe und mit Anweisungen für die Applikation auf die Aufgabe vorbereitet. Nach dem Lesen der Beschreibung wurde dem Teilnehmer zusätzlich noch ein kurzes Demonstrationsvideo [21] gezeigt. An diesem konnte der Teilnehmer sehen, wie das *Laut Denken*-Protokoll am Beispiel einer Webseite funktioniert. Dem Teilnehmer sollte durch dieses Video demonstriert werden, wie er an die Aufgabe herangehen soll und es sollte zudem ein besseres Verständnis geben, was genau von ihm verlangt wird.

Wenn der Teilnehmer Schwierigkeiten mit der Aufgabe hatte oder zu wenig seine Gedanken äußerte, versuchte der Versuchleiter durch Fragen den Teilnehmer zu animieren, seine Eindrücke und Gedanken bezüglich der Applikation zu schildern. Hatte der Teilnehmer die Aufgabe abgeschlossen, konnte er das Tablet ablegen und zum nächsten Schritt übergehen, welcher aus zwei Fragebögen bestand, die auf Papier auszufüllen waren.

Schritt 5 & 6: Vergleichsfragebogen und SUS-Fragebogen, Belohnung und Aufklärung

Der erste Fragebogen war ein Vergleichsfragebogen, welcher die Vorlieben der Teilnehmer zum Thema Darstellung der Fragen und Antwortmöglichkeiten abfragte. Hier musste der Teilnehmer entscheiden, welche Darstellungen für Fragen mit einer oder mehrere Antwortmöglichkeiten er bevorzugt. Dabei standen Antworten zum Ankreuzen oder eine grafische Darstellung der Antwortmöglichkeiten, wie er sie in der Aufgabe am mobilen Endgerät sehen konnte, zur Auswahl. Unter den grafischen Darstellungen der Antwortmöglichkeiten waren unter anderem eine Slidebar und eine Darstellung zur Angabe mehrerer Jahre als Antwortmöglichkeiten gegeben (Abb.4.2). Zusätzlich sollte der Teilnehmer angeben, ob er das Ausfüllen von Fragebögen auf Papier oder auf mobilen Endgeräten bevorzugt.

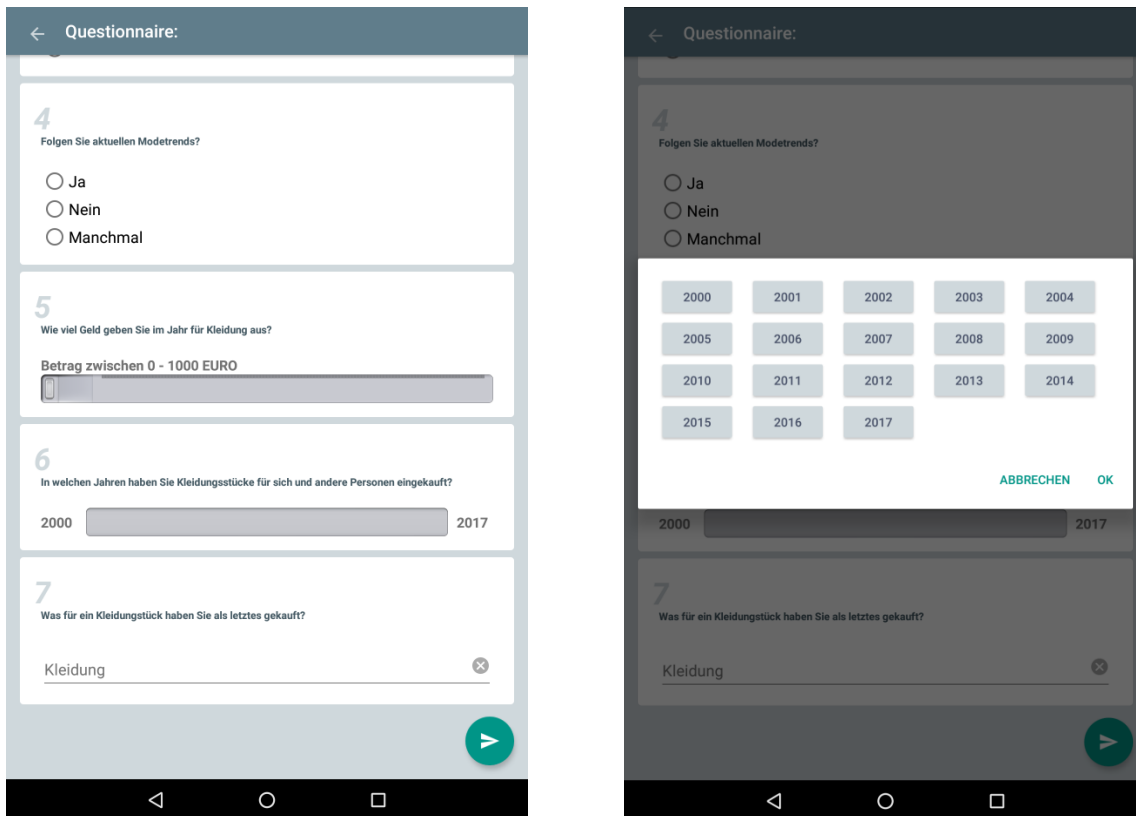


Abbildung 4.2: Darstellung von der Sidebar unter Frage 5 (links) und die Darstellung der Jahre als Antwortmöglichkeiten (rechts)

Darauf folgte ein standardisierter System Usability Scale - SUS Fragebogen zur allgemeinen Abfrage der Benutzbarkeit der Applikation, wo Aussagen zur Komplexität der Benutzbarkeit, Lernfähigkeit, Verständnis der Applikation, an einer Likertskala angekreuzt werden konnten.

4.3. Musterfragebogen

Als Unterstützung für die Benutzerstudie wurden Musterfragebögen zugezogen. Zwei der Fragebögen, die in dieser Benutzerstudie vorkamen, sind übliche Fragebögen einer Usability-Testung. Der erste ist der demografischer Fragebogen und der zweite der standardisierter Usability-Fragebogen.

4.3.1 Demographie Fragebogen

Mit Hilfe eines demografischen Fragebogens wurden allgemeine Informationen zu Geschlecht, Alter, Bildungsgrad abgefragt. Einige Fragen wurden hinzugefügt, um die Informationen, die für diese Studie zusätzlich relevant waren, abzufragen. Diese beziehen sich auf Besitz eines mobilen Endgeräts und dessen Benutzung, sowie Interesse und Kenntnis im Bereich mobiler Applikationen.

4.3.2 System Usability Scale (SUS)

Für die Bearbeitung der Aufgabe am mobilen Endgerät konnte keine genaue Zeit vorausgesagt werden. Diese wurde auf ungefähr 10 bis 15 Minuten geschätzt, mit der Berücksichtigung, dass sie von Teilnehmer zu Teilnehmer unterschiedlich ausfallen kann. Die benötigte Zeit hängt stark von den Aussagen des Teilnehmers ab, wie viel während der Aufgabe erzählt wurde und wie detailliert diese Aussagen waren. Da die Dauer von ca. 30 Minuten für einen Durchlauf nicht überschritten werden sollte, ist die Entscheidung, bei der Auswahl der standardisierten Usability-Fragebögen, auf den (System Usability Scale) SUS-Fragebogen [19] gefallen. Der kurze Fragebogen fasst die Usability-Merkmale gut zusammen und liefert, schon bei wenigen Teilnehmern, verwendbare Ergebnisse [20]. Dadurch zeichnete sich dieser Fragebogen, im Vergleich zu den anderen, als am besten geeignet für diese Benutzerstudie aus.

4.4. Lautes Denken

Die Methode *Lautes Denken* ermöglicht es durch lautes Äußern der Gedanken während einer Testung, Einblicke in die mentalen Prozesse der Versuchsperson zu erhalten. In der Softwareentwicklung wird die Methode in erster Linie zur Testung von Gebrauchtauglichkeit und Bedienungsfreundlichkeit von Systemen eingesetzt. Die Methode ermöglicht ebenfalls einen weiten Einblick über das Vorgehen der Teilnehmer beim Auftauchen eines Problems, während der Bearbeitung der Aufgabe. In kurzer Zeit können reichhaltige Informationen über relevante Aspekte der Benutzung des Systems gesammelt werden. Die Daten werden entweder durch Notizen vom Versuchsleiter, durch Videomitschnitte oder Logfiles

gesammelt. Durch Erklärung und konkrete Beispiele soll der Teilnehmer auf die Testung vorbereitet werden. Hat der Teilnehmer Probleme oder äußert sich zu wenig, kann der Versuchsleiter durch Fragen oder beispielhafte Aussagen ihm dabei helfen [25]. Die Auswertung der Aufnahmen und des Schriftmaterials muss in geeigneter Form aufgearbeitet werden. Die wichtigsten Aussagen aller Teilnehmer müssen schriftlich erfasst und sinnvoll interpretiert werden. Am Ende werden die Aussagen analysiert und Schlussfolgerungen und Vorschläge zur Verbesserung des Systems erarbeitet.

Die Methode bietet einige Vor- und Nachteile. Vorteile der Methode sind, dass sie günstig, robust, flexibel, einfach zu erlernen und überzeugend ist. Ergebnisse stehen unmittelbar zur Verfügung. Nachteile der Methode sind, dass es eine unnatürliche Situation darstellt, das Benutzerverhalten vorbelastet sein kann und, dass der Teilnehmer gefilterte Aussagen tätigt [21].

In der hier vorgestellten Benutzerstudie, wurden die Teilnehmer auf den Teil der Studie mit dem *Laut Denken*-Protokoll entsprechend vorbereitet. Neben der kurzen Beschreibung in der Aufgabenstellung, wurden die Teilnehmer noch einmal ausdrücklich darum gebeten während der Aufgabe laut ihre Meinung zu äußern. Das heißt, die Teilnehmer sollten die Applikation auf dem mobilen Endgerät testen und währenddessen sagen, was ihnen daran gefällt und was nicht, was hilfreich ist, usw. Um diese Vorstellung dem Teilnehmer näher zu bringen, wurde ein kurzes Demonstrationsvideo [21], in Dauer von einer Minute, gezeigt. In diesem Video wird die Webseitenpräsentation der Organisation *Ambassadors for Life* [23] angezeigt. Man hört eine Frau, die versucht auf der Seite herauszufinden, was die Organisation macht. Dabei klickt sie sich durch die einzelnen Menüpunkte und erzählt, was ihr als erstes ins Auge fällt, was für sie verständlich ist und was nicht, und was ihr an der Webseite gefällt.

Nachdem das Video den Teilnehmern gezeigt wurde, wurden sie gebeten mit der Aufgabe anzufangen und sich bei der Auswahl der Fragebögen in der Applikation, an die Anweisungen der Aufgabenstellung zu halten. Bevor der Versuchsleiter dem Teilnehmer das Tablet übergab, wurde die Aufnahme der Oberfläche des Tablets, durch die Applikation *AZ Screen Recorder* gestartet. Danach übergab er das Tablet dem Teilnehmer, der anschließend mit der Aufgabe beginnen konnte.

Die Aufgabe bestand darin, auf dem mobilen Endgerät die Applikation *Questionnaire* zu öffnen, das Menü mit den Fragebögen zu finden und die Fragebögen *Mode* und *Elektronik*

auszufüllen. Von den beiden Fragebögen gab es jeweils zwei Versionen. Eine Version wurde mit Antwortmöglichkeiten zum Ankreuzen entwickelt, während die andere Version mit unterschiedlichen grafischen Darstellungen der Antwortmöglichkeiten, wie freie Texteingabe, Slider, Range entwickelt wurde. Jedem Teilnehmer wurden von jeder Version ein solcher Fragebogen zugeteilt. Welche Fragebögen er bekommt, wurde zufällig entschieden. War der Teilnehmer mit beiden Versionen der Fragebögen fertig, folgte passend zu der Aufgabe ein Vergleichsfragebogen.

4.5. Vergleich der Design-Elemente in Fragebögen

In der Aufgabe mit dem *Laut Denken*-Protokoll, gab es für die zwei Fragebogen in der Applikation jeweils zwei verschiedene Versionen. In beiden Versionen kamen die selben Fragen vor. Die Versionen unterschieden sich lediglich bei der Darstellung der Antwortmöglichkeiten. Was bei einem als Frage mit Antwortmöglichkeiten zum Ankreuzen angezeigt wurde, war beim anderen mit grafischer Darstellung der Antwortmöglichkeiten gegeben und umgekehrt. So hatte die erste Hälfte der Teilnehmer die Versionen *Mode 1 und Elektronik 2 Fragebogen* und die zweite Hälfte *Mode 2 und Elektronik 1 Fragebogen* (Tabelle 1). Dieses Vorgehen sollte eine Randomisierung [27] bei der Studie ermöglichen und den Einfluss der Reihenfolge der Fragebögen auf die Ergebnisse der Studie ausschließen.

	Teilnehmer 1-15	Teilnehmer 16-30
Darstellungen der Fragebogen	Mode 1	Mode 2
	Elektronik 2	Elektronik 1

Tabelle 1: Aufteilung der Versionen der Fragebögen auf die Teilnehmer

Im Anschluß der abgeschlossenen Aufgabe am mobilen Endgerät, wurde dem Teilnehmer ein Vergleichsfragebogen vorgelegt. Im Fragebogen wurde die Meinung des Teilnehmers über die verschiedenen Darstellungen der Antwortmöglichkeiten abgefragt. Der Fragebogen bestand aus fünf Fragen mit Antwortmöglichkeiten zum Ankreuzen. Außer zur Darstellung, wurde auch abgefragt, ob der Teilnehmer das Ausfüllen der Fragebögen, auf Papier oder auf mobilen Endgeräten priorisiert. Die Auswertung des Vergleichsfragebogens sollte dabei helfen, die Hypothesen zu überprüfen.

5

Evaluierung

Dieses Kapitel stellt die Ergebnisse der Benutzerstudie vor. Diese gliedern sich in die subjektiven Meinungen der Teilnehmer bei dem *Laut Denken*-Protokoll, Vergleich der Design-Elemente und Auswertung des standardisierten Usability-Fragebogen zur allgemeinen Benutzbarkeit.

5.1. Demographische Daten

Bei der Studie nahmen insgesamt 30 Teilnehmer teil, 22 weibliche und 8 männliche. Das Durchschnittsalter lag bei 25 Jahren bei einer Schwankungsbreite von 19 bis 36 Jahren. Über die Hälfte der Teilnehmer (56,67 %) waren Studenten. [28] diskutiert dass Studenten als angemessener Ersatz für Fachexperten in empirischen Studien in Frage kommen. Es wurden keine besonderen Fähigkeiten oder Wissen von den Teilnehmern erwartet.

Ein Tablet besaßen lediglich 23,33 % der Teilnehmer, im Vergleich zu einem Smartphone, bei dem alle Teilnehmer angaben, ein Smartphone zu besitzen. Die Teilnehmer sollten angeben, welches Betriebssystem ihr Smartphone hat. Daraus ergab sich folgendes Diagramm:

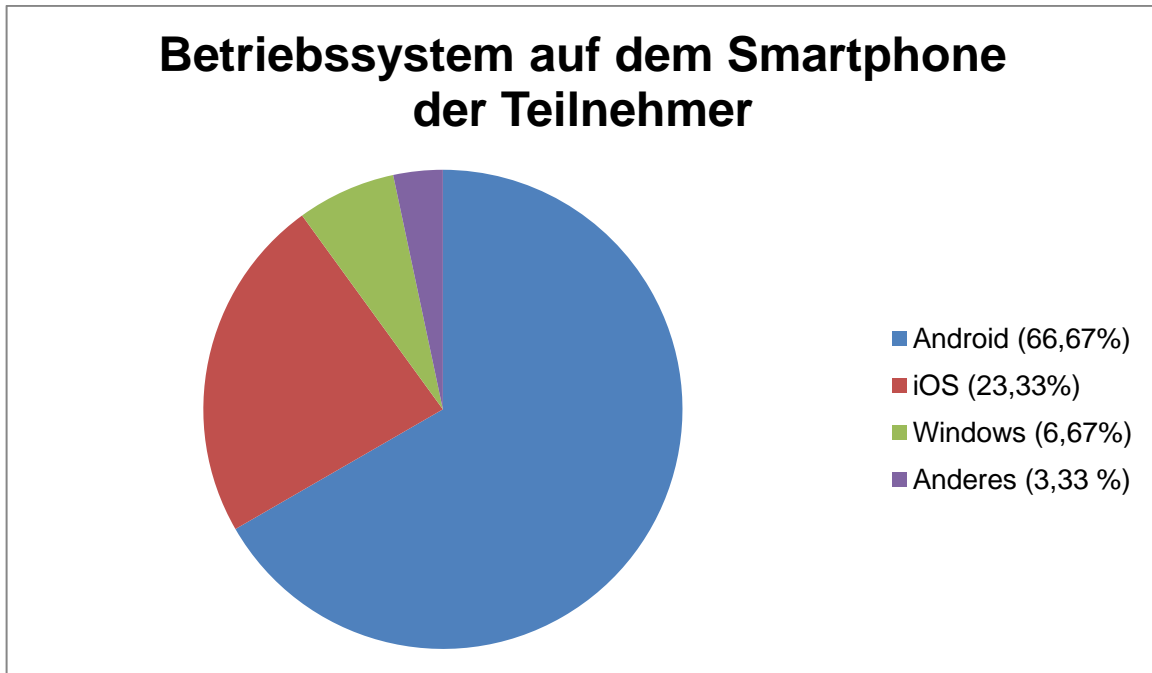


Abbildung 5.1: Betriebssystem auf dem Smartphone der Teilnehmer

Der Großteil der Teilnehmer (86,67 %) gab an oft oder sehr oft ihr mobiles Endgerät zu benutzen. 9 der 30 Teilnehmer gaben an Applikationen zur Datensammlung benutzt zu haben.

5.2 Ergebnisse

Im folgenden Abschnitt werden die Ergebnisse der einzelnen Aufgaben der Benutzerstudie vorgestellt. Als erstes werden die Aspekte während des Ablaufs und die Ergebnisse des *Laut Denken*-Protokolls beschrieben und daraus hergeleitete Verbesserungsvorschläge aufgelistet. Danach folgend die Ergebnisse des Vergleichsfragebogens zum Vergleich der Design-Elemente in der Applikation und schließlich die Auswertung des standardisierten SUS-Fragebogens.

5.2.1 Lautes Denken-Methode

Die Aufgabe am mobilen Endgerät war mit dem *Laut Denken*-Protokoll verbunden. Der Teilnehmer sollte während der ganzen Bearbeitung der Aufgabe seine Gedanken laut

äußern und seine Eindrücke über die Applikation schildern. Zweck dieser Aufgabe war es die Benutzbarkeit der Applikation zu testen. Dabei sollte auch das Design bewertet und mögliche Verbesserungsvorschläge hergeleitet werden.

Während der Bearbeitung der Aufgabe wurden Oberfläche des Tablets und die gesprochenen Kommentare aufgenommen. Das Material diente zur Auswertung des *Laut Denken*-Protokolls. Die durchschnittliche Zeit zur Bearbeitung der gesamten Aufgabe betrug 7:30 Minuten, wobei verschiedene Teilnehmer zwischen 5 und 15 Minuten benötigten. Jeder Teilnehmer der Studie schloss die gesamte Aufgabe am mobilen Endgerät erfolgreich ab. Jede Antwort in den Fragebögen wurde beantwortet und jeder Teilnehmer äußerte sich währenddessen zu der Benutzung und Darstellung der Applikation. Es sind keine kritischen Fehler während der Bearbeitung aufgetreten. Das mobile Endgerät und die Applikation funktionierten einwandfrei.

Während der Bearbeitung der Aufgabe sind jedoch kleinere, nicht kritische Fehler aufgefallen. Einer davon war, dass der Weiter-Button auf der letzten Seite die Teilnehmer verwirrte, da beim Drücken des Buttons nichts passiert ist. Der Weiter-Button auf der letzten Seite dient eigentlich dazu, die Teilnehmer abzufragen, ob ihre Antworten gespeichert werden sollen. Da die eingegebenen Daten der Teilnehmer im Fragebogen nicht relevant für die Benutzerstudie waren und nicht gespeichert wurden, hatte der Weiter-Button auf der letzten Seite keine Funktion. Außerdem hat die Auflistung aller Fragebögen im Menü die Teilnehmer irritiert, da diese nicht nummeriert waren und sie selber nachschauen mussten, welchen Fragebogen sie auswählen müssen. Diese kleineren Fehler haben jedoch nicht das erfolgreiche Abschliessen der Aufgabe verhindert. Einige Teilnehmer haben sehr ausführlich ihre Eindrücke über die Applikation geschildert, wohingegen einige nur auf die wesentliche Merkmale der Applikation eingegangen sind. Während der Auswertung der Aufnahmen wurde festgestellt, dass es einige Punkte gab, die fast alle Teilnehmer angesprochen haben. Diese Punkte werden in folgender Tabelle unterteilt und die positiven und negativen Kommentare dazu werden aufgezählt.

• Design	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ Organisiert + Keine Ablenkung durch bunte Farben + Seiten nicht überladen	- Viel leeren Platz auf der Seite - Design zu schlicht - Schriftgröße bei den Fragen etwas zu klein
• Funktionen	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ Sprachauswahl	- Auf der letzten Seite fehlt ein Feld für Feedback oder persönliche Anmerkungen
• Fragebögen	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ sehr übersichtlich + Überschriften und Einleitungstext sehr hilfreich + Fragebogen, indem mehr Single/Multiple Choice Fragen waren, kam besser an	- Einleitungstext überschneidet sich mit der Überschrift - Antwortfelder bei Single/Multiple-Choice Fragen zu nah beieinander - Doppelte Überschrift „Danke“ auf der letzten Seite - Bei der Aufzählung aller Fragebögen im Menü fehlt eine Nummerierung
• Sidebar	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ intuitiv	- keine Skalierung
• Range	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ Range Darstellung bei Fragen mit mehreren Antwortmöglichkeiten, kam besser an als Multiple Choice Darstellung	- Es fehlte eine Möglichkeit, um alle Antworten auszuwählen - Nicht sofort ersichtlich, dass man die Leiste drücken muss, damit das Fenster mit den Antworten aufgeht
• Weiter-Button	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ befindet sich immer an der selben Stelle	- Sollte erst auftauchen, wenn alle Fragen beantwortet sind - Weiter-Button auf der letzten Seite hatte keine Funktion und hat die Teilnehmer verwirrt
• Fragen	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ Möglich die Antwort zu ändern, falls man sich vertippt	- Tastatur überdeckt letzte Frage - Antwortmöglichkeiten vorgeben anstatt freier Texteingabe - Bei Multiple-Choice Fragen fehlte ein Button, um alle Antworten auszuwählen
• Bedienbarkeit	
<i>Positive Kommentare:</i>	<i>Negative Kommentare:</i>
+ einfach + Fragebögen sind einfach zu starten	

Tabelle 2: Wichtige Punkte die beim Laut Denken-Protokoll angemerkt wurden

Verbesserungsvorschläge:

Die Teilnehmer haben sich unterschiedlich zum Design der Applikation geäußert. Einige fanden, dass es durch das schlichte Design nicht vom ablenkt und die Konzentration auf die Fragebögen lenkt. Andere hingegen fanden das Design zu schlicht und hätten sich mehr Farben oder sogar Bilder gewünscht. Hier ist es Geschmackssache, ob einem das Design gefällt oder nicht. Jedoch sind bei psychologischen oder medizinischen Fragebögen solche Ablenkungen durch Farben und Bilder nicht vorteilhaft. Die Gestaltung des Fragebogens ist vom Anwendungsszenario abhängig.

Die Sprachauswahl kam bei den Teilnehmern sehr gut an. Als zusätzliche Funktion könnte ein Feld auf der letzten Seite eingefügt werden, für Feedback und persönliche Anmerkungen, was im Konfigurator erstellt werden kann.

Die Fragebögen in der Applikation waren übersichtlich dargestellt und die Überschriften und Einleitungstext fanden die Teilnehmer sehr hilfreich, um zu erfahren, worum es in dem Fragebogen geht. Beim Öffnen des Fragebogens haben sehr viele Teilnehmer angemerkt, dass sich der Text mit der Überschrift überschneidet und sie das optisch nicht ansprechend finden. Einige dachten, dass es ein Fehler bei der Darstellung ist. Hier wäre es ratsam noch einmal das Design zu überdenken oder einen größeren Abstand zwischen Überschrift und Text einzufügen, damit sich diese nicht überschneiden.

Es wurde zusätzlich angemerkt, dass die Antwortfelder bei Single-Choice / Multiple-Choice-Fragen zu nah beieinander liegen. Teilnehmer mit etwas größeren Fingern fanden das ärgerlich, da sie sich beim Auswählen der Antwort vertippt haben.

Das Design-Elemente Slider kam bei den Teilnehmern gut an. Es war für sie intuitiv bedienbar. Der einzige Nachteil war, dass es keine Skalierung gab, heißt auf dem Slider waren keine Werte vorgegeben und die Teilnehmer mussten schätzen bei welchem Wert sie sich befinden. In der Zukunft sollte beachtet werden, dass bei der Verwendung des Sliders als Antwortmöglichkeit, immer eine Skalierung dabei ist, um das Beantworten der Frage einfacher zu gestalten.

Das Design-Element Range kam im Vergleich zu Multiple-Choice Darstellung, besser bei den Teilnehmern an. Sie fanden die grafische Darstellung schöner. Am Anfang war für viele der Teilnehmer nicht sofort ersichtlich, wie man die Frage mit dem Range-Element beantworten soll. Auf die Lösung kamen sie, indem sie Verschiedenes ausprobiert haben, wobei sie mehr oder wenig zufällig auf die Leiste gedrückt haben. Die Teilnehmer hätten sich hier einen deutlichen Hinweis gewünscht. Eine mögliche Lösung ist in der Abbildung 5.2 dargestellt.

The image shows a survey question interface. At the top left, there is a large number '6'. Below it, the question text reads: 'In welchen Jahren haben Sie Kleidungsstücke für sich und andere Personen eingekauft?'. Underneath the question, there is a smaller instruction: 'Drücken Sie die Leiste, um das Fenster mit den Antwortmöglichkeiten zu öffnen'. Below this instruction is a horizontal range slider. The slider has a light gray track with rounded ends. On the left end of the track, the year '2000' is written, and on the right end, the year '2017' is written. The slider itself is a darker gray bar that spans the entire width of the track between the two years. Below the question, there is a large number '7'.

Abbildung 5.2: Mögliche Lösung für das Problem beim Range-Element

Der Weiter-Button sollte erst auftauchen, wenn alle Fragen beantwortet sind. Durch die ständige Präsenz des Weiter-Button, sind die Teilnehmer verwirrt. Bei einem Fragebogen haben einige Teilnehmer übersehen, dass es unten auf der Seite weitergeht. Sie haben infolgedessen den Weiter-Button geklickt und wobei nichts passiert ist. Schließlich haben die Teilnehmer bemerkt, dass sie den Fragebogen nicht vollständig ausgefüllt haben und es weiter unten auf der Seite weitergeht.

Weit definierbare Begriffe wie Häufigkeit oder andere offene Fragen sollten lieber mit vorgegebenen Antwortmöglichkeiten gegeben sein, sonst sind die Teilnehmer unsicher, wie sie die Frage beantworten sollen, was sich auf die Zufriedenheit der Teilnehmer auswirken kann.

Bei Fragen mit mehreren Antwortmöglichkeiten wäre es sinnvoll einen Button hinzuzufügen, mit dem alle Antworten ausgewählt werden können. Alle Antworten einzeln auszuwählen, empfinden die Teilnehmer als mühselig, was sich ebenfalls negativ auf die Zufriedenheit der Teilnehmer auswirken kann.

Allgemein waren die Teilnehmer zufrieden mit der Bedienung der Applikation und haben keine Hilfestellungen benötigt. Fast alle Funktionen waren intuitiv und einfach bedienbar. Die Applikation machte insgesamt einen guten Eindruck auf die Studienteilnehmer.

5.2.2 Vergleichsfragebogen

In diesem Abschnitt werden die Ergebnisse des Vergleichsfragebogens vorgestellt, die dazu dienen die Einstellungen der Teilnehmer gegenüber einfachen Darstellungen der Antwortmöglichkeiten und neuen Design-Elementen zu vergleichen.

Der Vergleichsfragebogen bestand aus fünf Fragen, die speziell auf den Vergleich der verschiedenen Darstellungen der Antwortmöglichkeiten eingehen. Da in der Applikation neue Design-Elemente hinzugefügt wurden, sollten diese mit den einfachen Darstellungen (Single-Choice / Multiple-Choice) verglichen werden. Damit die Teilnehmer eine Vorstellung haben, wie diese Darstellungen aussehen, wurden einfache, sowie neue Design-Elemente in der Aufgabe am mobilen Endgerät verwendet.

Die Abbildung 5.3 zeigt die Anzahl der Teilnehmer bei Beantwortung der ersten Frage, welche die jeweiligen Antworten ausgewählt haben. Dabei bevorzugten 63,33% der Teilnehmer das Ankreuzen der Antworten (Single-Choice / Multiple-Choice Darstellungen) bei Fragebögen auf dem mobilen Endgerät.

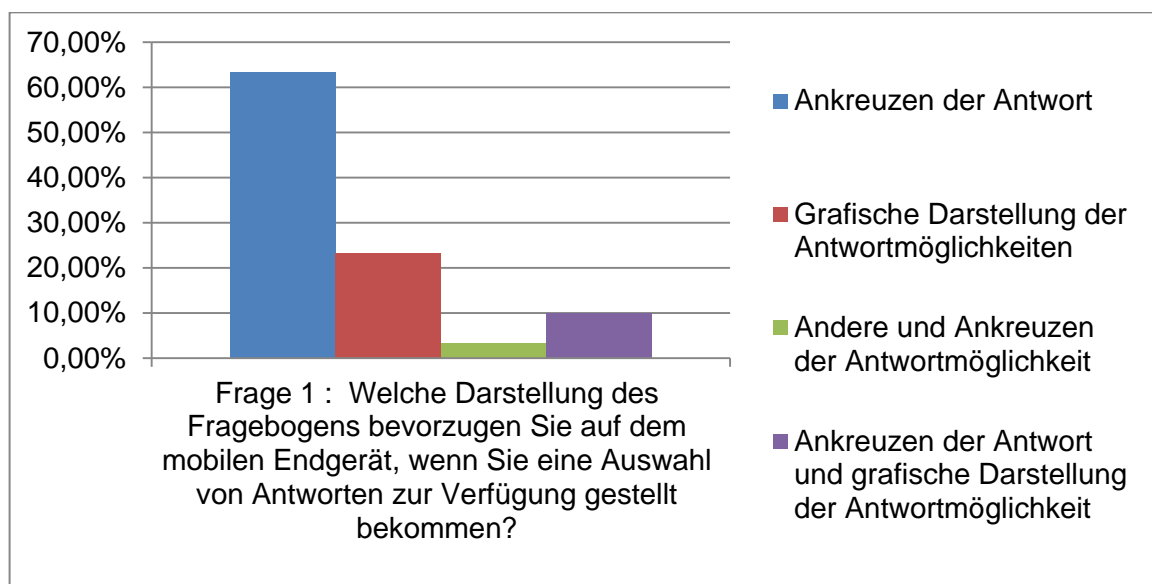


Abbildung 5.3: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 1 des Vergleichsfragebogens

Bei der zweiten Frage liegt ein eindeutige Präferenz vor (Abb.5.4), in der 70 % der Teilnehmer, für Fragen mit *einer* möglichen Antwort, die Antwortmöglichkeiten zum Ankreuzen auswählen und 30 % eine grafische Darstellung der Antwortmöglichkeiten bevorzugen.

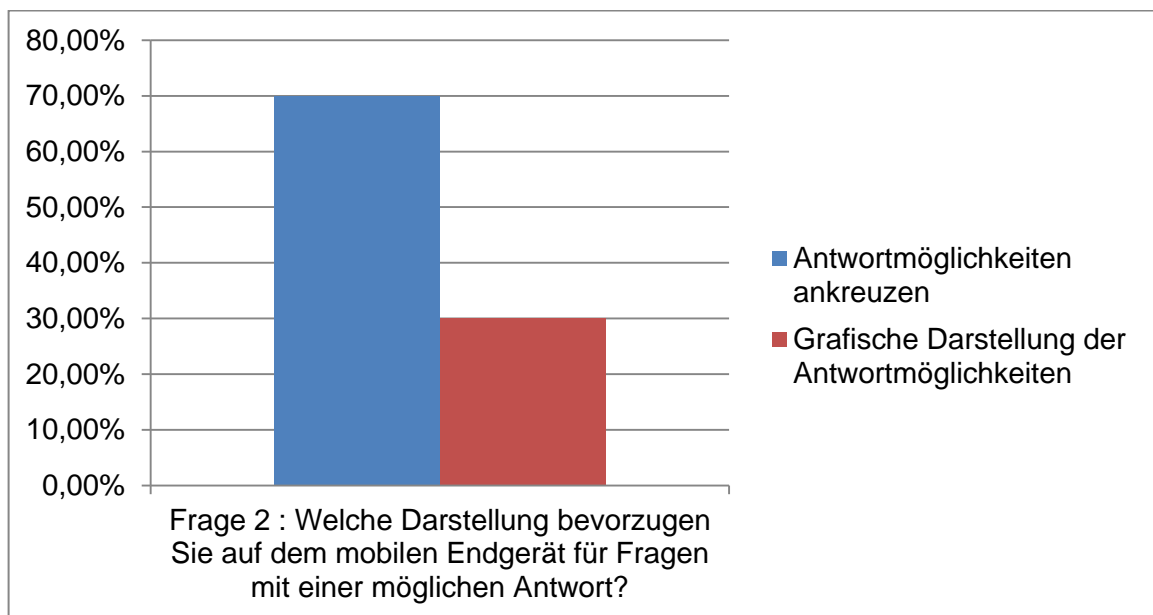


Abbildung 5.4: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 2 des Vergleichsfragebogens

In Frage 3 wurde das gleiche Verhältnis abgefragt, lediglich mit dem Unterschied, dass es sich hier um Fragen mit *mehreren* Antwortmöglichkeit handelt. In Abbildung 5.5 ist, im Vergleich zur letzten Frage, ein Unterschied zu sehen, indem 63,33 % Antwortmöglichkeiten zum Ankreuzen wählen und 36,67 % die grafische Darstellung der Antwortmöglichkeiten.

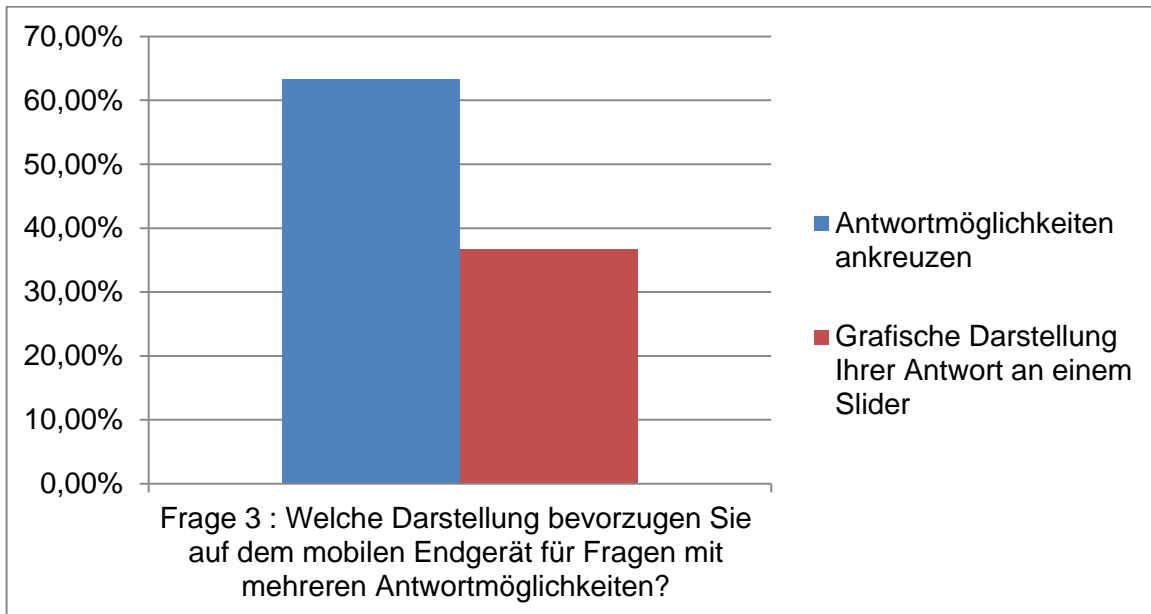


Abbildung 5.5: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 3 des Vergleichsfragebogens

Bei der Frage zur einfacheren Bedienbarkeit der zwei verschiedenen Darstellungen der Antwortmöglichkeiten, haben sich 80 % der Teilnehmer für Fragebogen mit Antwortmöglichkeiten zum Ankreuzen ausgesprochen und lediglich 16,67 % für Fragebogen mit verschiedenen grafischen Darstellungen und 3,33 % fanden beide Darstellung einfach zu bedienen.

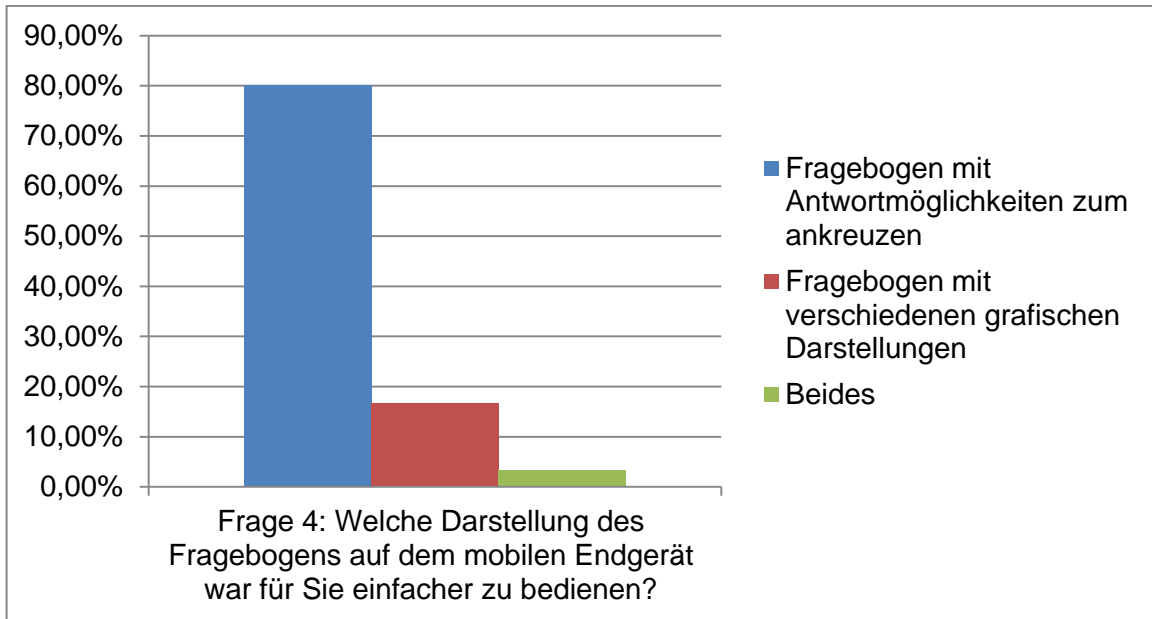


Abbildung 5.6: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 4 des Vergleichsfrageboges

Bei der Frage zum Ausfüllen von Fragebögen auf Papier oder mobilen Endgerät, konnte kein eindeutiges Ergebnis hervorgebracht werden. 50 % der Teilnehmern wurden das Ausfüllen am mobilen Endgerät bevorzugen, 43,33 % das Ausfüllen auf Papier und 6,67 % haben beide Möglichkeiten ausgewählt. Somit kann hier keine eindeutige Präferenz angegeben werden, sondern die Wahl hängt von persönlichen Vorlieben des Teilnehmers ab.

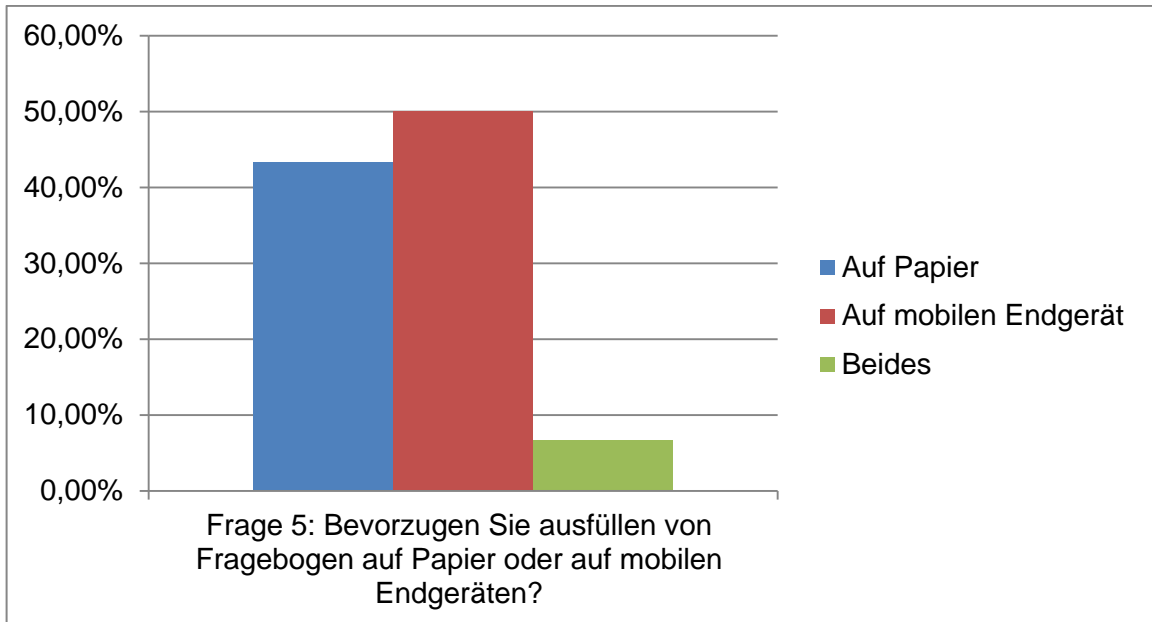


Abbildung 5.7: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 5 des Vergleichsfragebogens

Die Ergebnisse zeigen, dass die Teilnehmer eher zu Antwortmöglichkeiten zum Ankreuzen tendieren. Ein möglicher Grund dafür könnte sein, dass die Teilnehmer die einfachere Variante wählen. Bei Antwortmöglichkeiten zum Ankreuzen, sind die Teilnehmer schneller mit dem Ausfüllen fertig. Durch die vorgegebenen Antworten müssen sie nicht lange über eine Antwort nachdenken, während bei grafischen Darstellungen der Antwortmöglichkeiten, sie sich erstmal eine Antwort überlegen müssen oder länger zum Beantworten der Frage brauchen.

5.2.3 Auswertung des SUS-Fragebogen

Der SUS-Fragebogen [19] besteht aus 10 Aussagen. Die Aussagen 1, 3, 5, 7 und 9 stellen eine positive Aussage dar und werden aufsteigend auf der Likertskala mit „Stimme gar nicht zu“ bis „Stimme voll zu“ bewertet. Umgekehrt verhält es sich bei den Aussagen mit einer geraden Zahl davor. Ist der Fragebogen vollständig ausgefüllt, werden die sogenannten SUS-Scores der Aussagen addiert und mit 2,5 multipliziert. Die Scores der einzelnen Aussagen sind für sich nicht aussagekräftig [19]. Es folgt ein insgesamt SUS-Score zwischen 0 und 100. Dieser kann als Prozentwert interpretiert werden, wobei Werte zwischen 90-100 % einem nahezu perfektem System entsprechen, Werte zwischen

70-90 % einem guten bis sehr guten System und alle Werte unter 60 % Mängel am System im Bezug auf die Benutzerfreundlichkeit vorweisen (Abb.5.8).

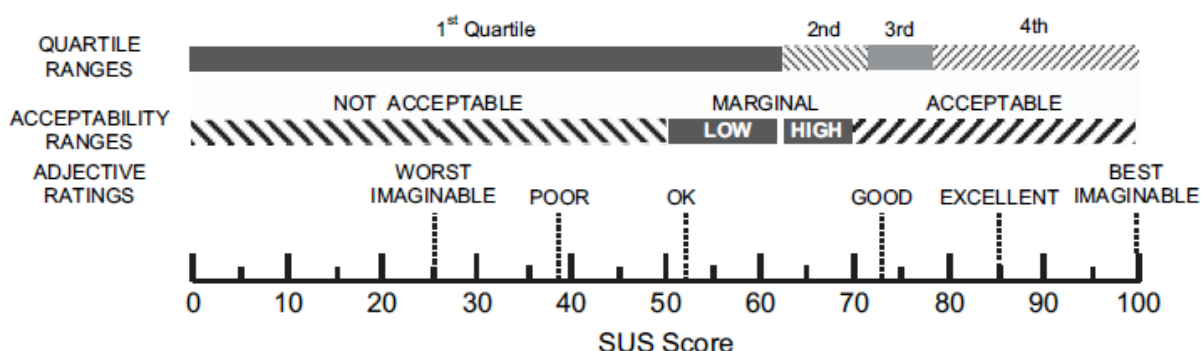


Abbildung 5.8: Bangor Adjective Scale [29]

Die Ergebnisse der SUS-Scores von jedem einzelnen Teilnehmer können der Abbildung 5.9 entnommen werden, wobei der höchste Wert bei 95 und der niedrigste Wert bei 47,5 liegt. Die Auswertung aller SUS-Scores hat einen Durchschnittswert von 77 % ergeben. Daraus folgt, dass das System, im diesem Fall die Applikation, laut [29] mit gut bis sehr gut abschneidet.

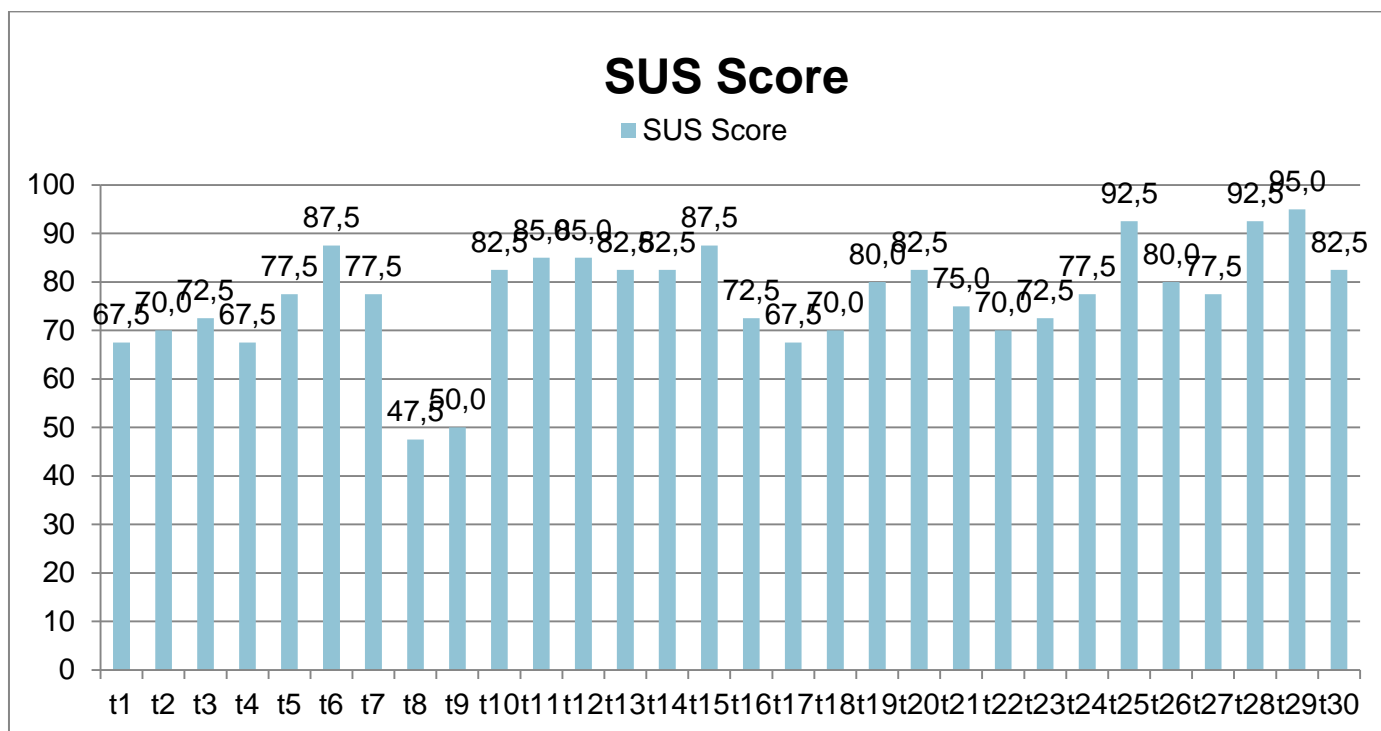


Abbildung 5.9: Ergebnisse des SUS-Scores der einzelnen Teilnehmer

5.3. Auswertung

Die Auswertung der vorgestellten Ergebnisse der Benutzerstudie, dienten auch zur Überprüfung der aufgestellten Hypothesen.

Bei der Anzahl von 30 Teilnehmer der Studie, wurden diese in zwei Gruppen aufgeteilt. Die erste Gruppe der Teilnehmer (Teilnehmer 1-15) bekamen bei der Aufgabe am mobilen Endgerät die Fragebögen-Versionen *Mode 1* und *Elektronik 2*. Die zweite Gruppe der Teilnehmer (Teilnehmer 16-30) bekamen die Fragebögen-Versionen *Mode 2* und *Elektronik 1*. Die Versionen der Fragebögen unterschieden sich bei der Darstellung der Antwortmöglichkeiten, so dass in der einen Version mehr Antwortmöglichkeiten zum Ankreuzen und in der zweiten mehr neue Design-Elemente zur Darstellung der Antwortmöglichkeiten benutzt wurden. Unter dieser Bedingung wurden die Ergebnisse des Vergleichsfragebogens verglichen und der Mittelwert ausgerechnet. Anschließend wurde ein T-Test gerechnet, woraus sich folgende Ergebnisse ergaben:

In der ersten Hypothese der Studie H0a wurde angenommen, dass die neuen Design-Elemente gleich gut sind wie die bekannten Antwortmöglichkeiten zum Ankreuzen sind. Anhand der Ergebnisse konnte festgestellt werden, dass $(T(\text{Freiheitsgrade})=29, p<.5)$ der p-Wert signifikant abweichend von 0.5 ist, wodurch die Nullhypothese H0a verworfen wird.

Das gleiche zeichnete sich für die Nullhypothese H1a ab. Hier wurde angenommen, dass das Ausfüllen von Fragebögen mit neuen Design-Elementen genauso einfach ist, wie das Ausfüllen von Fragebögen mit Antwortmöglichkeiten zum Ankreuzen. Anhand der Ergebnisse $(T(\text{Freiheitsgrade})=29, p<.5)$, wurde auch diese Nullhypothese H1a verworfen.

Somit werden die Alternativhypothesen H0b und H1b angenommen.

Die H0b hat aber nach auswerten des Vergleichsfragebogens ergeben, dass die Teilnehmer die Antwortmöglichkeiten zum Ankreuzen bevorzugen und nicht die grafische Darstellung der Antwortmöglichekeiten, womit die H0b ebenfalls nicht bestätigt werden kann.

H1b ist eindeutig bestätigt mit der Aussage, dass die Teilnehmer es als einfacher empfinden Fragebögen mit Antwortmöglichkeiten zum Ankreuzen auszufüllen, als die mit verschiedenen Design-Elementen wie Slider, freies Textfeld oder Range.

Bei der Testung der Nullhypothese H2a war der p-Wert größer als 0.5 und somit nicht signifikant. ($T(\text{Freiheitsgrade})=29, p>.5$), wodurch die Hypothese nicht verworfen wird.

Daraus ergibt sich auch, dass die Alternativhypothesen H2b verworfen wird und das Ausfüllen von Fragebögen am mobilen Endgerät nicht bevorzugt wird, im Vergleich zu Ausfüllen auf Papier.

6

Zusammenfassung und Ausblick

Ziel dieser Bachelorarbeit war die Evaluierung der Benutzbarkeit einer mobilen Anwendung anhand einer Studie. Bei der Studie haben insgesamt 30 Teilnehmer teilgenommen. In der Arbeit wurden einige kleinere Usability-Probleme der zur evaluierenden Anwendung aufgedeckt und Verbesserungsvorschläge erarbeitet. Gleichzeitig wurde festgestellt, dass die Teilnehmer der Studie mit der mobilen Anwendung grundsätzlich gut zurechtkamen und zufrieden waren. Dies ist aus dem Teil der erfolgreich abgeschlossenen Aufgaben und aus dem direkten Feedback der Teilnehmer abzulesen. Durch die Benutzerstudie konnten zudem auch neue Erkenntnisse zur Darstellung und Design der mobilen Anwendung gewonnen werden. Die Ergebnisse des Laut Denken-Protokolls haben fehlende Funktionen, wie ein Feld für Feedback und persönliche Anmerkungen oder einer Möglichkeit der Auswahl aller Antwortmöglichkeiten bei einer Frage aufgedeckt. Diese Funktionen lassen sich durch den Konfigurator einfügen. Außerdem wurden Verbesserungsvorschläge für die neuen Design-Elemente, wie Slider oder Range genannt. Die Teilnehmer waren der grafischen Darstellung der Antwortmöglichkeiten nicht abgeneigt, dennoch hat die Mehrheit der Teilnehmer die einfachere, bereits bekannte Möglichkeit zum Ausfüllen der Fragebögen bevorzugt. Aus den Ergebnissen und dem direkten Feedback der Teilnehmer geht als Grund für die Entscheidung hervor, dass die Teilnehmer deutlich schneller beim Ausfüllen des Fragebogens mit mehr Single-Choice / Multiple-Choice Fragen waren und die schon vorgegebenen Antworten ihnen beim Ausfüllen geholfen haben. Freie Texteingaben als Antwortmöglichkeiten kamen dagegenn bei den Teilnehmern nicht gut an.

Eine knappe Mehrheit der Teilnehmer hat das Ausfüllen der Fragebögen auf dem mobilen Endgerät anstatt auf Papier bevorzugt. Das spricht dafür, dass sich die Nutzer immer mehr an die digitale Version des Ausfüllens von Fragebögen gewöhnen.

6.1 Ausblick

Die mobile Applikation *Questionnaire* wurde anhand einer Benutzerstudie evaluiert. Die durchgeführte Studie könnte als Beispiel für zukünftige Testung der Usability von mobilen Anwendungen genutzt werden. Die Beschreibungen und der Vergleich der verschiedenen Methoden kann als Quelle für Ideen zur Evaluierung einer mobilen Anwendungen dienen.

Nach Aufzählung der Verbesserungsvorschläge für die getestete Applikation in Kapitel fünf, werden folgende Punkte als mögliche Weiterentwicklung der Applikation genannt:

1. Das Layout der Applikation ist schlicht und organisiert. Sinnvoll wäre es, die Schriftgröße etwas zu vergrößern. Das Design der Überschrift und des Einleitungstextes bei den Fragebögen könnte so geändert werden, dass sich die zwei Elemente nicht überschneiden. So könnten Verwirrungen und Fehlinterpretationen des Designs durch die Nutzer vermieden werden.
2. Im Hinblick auf eine Erweiterung der Applikation, wäre das Hinzufügen eines Feldes für Feedback und persönliche Anmerkungen am Ende der Fragebögen von Vorteil. Dies lässt sich im Konfigurator einfügen. Somit könnten die Nutzer auch Vorschläge zur Optimierung der Applikation machen.
3. Im eigentlichen Menü der Applikation wäre eine Unterteilung in Kategorien beim Menüpunkt *Fragebögen* von Vorteil. Aufteilung der Fragebögen nach Themen oder der Anwendungsszenarien würde die Suche nach einem bestimmten Fragebogen weiter erleichtern. Werden in der Applikation viele Fragebögen zur Verfügung gestellt, wäre das Einfügen einer kleinen Suchfunktion im Menü auch eine Möglichkeit der Erweiterung.
4. Beim eigentlichen Fragebogen ist es möglich, im Konfigurator verschiedene Optionen auszuwählen von Aufteilung der Fragen auf verschiedene Seiten bis

Einfügen von Bildern. Hier bieten sich ebenfalls verschiedene Möglichkeiten zur Darstellung der Antwortmöglichkeiten an. Bei Fragen mit mehreren Antwortmöglichkeiten, könnte zusätzlich eine Möglichkeiten zum Auswählen aller Antworten eingefügt werden. Nutzer, die alle Antworten auswählen wollen, würden durch die Option es wesentlich leichter haben. Im Gegensatz könnte es passieren, dass die Nutzer es zu umständlich finden alle Antworten einzeln auszuwählen und dabei die Frage nicht wahrheitsgemäß beantworten.

5. Die neu erstellten Design-Elemente erfordern ebenfalls kleine Verbesserungen. Bei dem Slider sollten die Werte auf dem Slider mitangezeigt werden, um das Beantworten der Frage zu erleichtern. Bei dem Design-Element Range sollte ein ausdrücklicher Hinweis zum Auswählen der Elemente hinzugefügt werden.

Die derzeitige mobile Applikation unterstützt das mobile Betriebssystem Android. Da viele Nutzer auch das Betriebssystem iOS auf ihrem mobilen Endgerät haben, wäre es sinnvoll die Applikation auch für das iOS Betriebssystem anzubieten.

Literaturverzeichnis

- [1] Balzert, H. (2010). *Lehrbuch der Softwaretechnik: Basiskonzepte und Requirements Engineering*. Springer-Verlag.
- [2] ISO (2011), *ISO/IEC 25010:2011, Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (SQuaRE) – System and Software Quality Models*, International Organisation for Standardisation, Geneva.
- [3] Herzwurm, G., & Mikusz, M. (2010). *Qualitätsmerkmale von Software. Kurbel ua (Hrsg., 2010)*. Version: 12.12.2016
- [4] Thielsch, M. T. (2008). *Inhalt, Usability und Ästhetik in der Bewertung durch Webnutzer*. In *Mensch & Computer* (pp. 441-444).
- [5] Dröschel, W., & Wiemers, M. (Eds.). (1999). *Das V-Modell 97: der Standard für die Entwicklung von IT-Systemen mit Anleitung für den Praxiseinsatz*. Walter de Gruyter GmbH & Co KG.
- [6] Fink, A. (2003). *The Survey Handbook* (Vol. 1). Sage.
- [7] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering*. Springer Science & Business Media.
- [8] Büttner, G. (2008). *Fragebögen und Ratingskalen*. Handbuch der Psychologie, Göttingen ua, 282-290.
- [9] Satterlee, E., McCullough, L., Dawson, M., & Cheung, K. (2015). *Paper-To-Mobile Data Collection: A Manual*. Version: 03.2015
- [10] Hildebrandt, A., Luthiger, J., Stamm, C., & Yereaztian, C. (2012). *Eine Kategorisierung mobiler Applikationen. IMVS Fokus Report*.
- [11] Runeson, P., Host, M., Rainer, A., & Regnell, B. (2012). *Case Study Research in Software Engineering: Guidelines and Examples*. John Wiley & Sons.
- [12] Herczeg, M. (2005). *Software-Ergonomie: Grundlagen der Mensch-Computer-Kommunikation*. Oldenbourg Verlag.
- [13] Nielsen, J., & Loranger, H. (2006). *Prioritizing Web Usability*. Pearson Education.
- [14] Figl, K., & Mensch, U. F. I. V. T. (2009). *ISONORM 9241/10 und Isometrics: Usability-Fragebögen im Vergleich*. In *Mensch & Computer* (Vol. 9, pp. 143-152).
- [15] Nielsen, J. (1994). *Usability Engineering*. Elsevier.

- [16] Schobel, J., Pryss, R., Wipp, W., Schickler, M., & Reichert, M. (2016). *A Mobile Service Engine Enabling Complex Data Collection Applications*. In: *14th International Conference on Service-Oriented Computing (ICSOC 2016)*, October 10-13, 2016, Banff, Alberta, Canada.
- [17] Albert, W., & Tullis, T. (2013). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Newnes.
- [18] Lethbridge, T. C., Sim, S. E., & Singer, J. (2005). *Studying Software Engineers: Data Collection Techniques for Software Field Studies*. *Empirical Software Engineering*, 10(3), 311-341. Springer
- [19] Brooke, J. (1996). *SUS - A quick and dirty Usability Scale*. *Usability Evaluation in Industry*, 189 (194), 4-7. London, United Kingdom
- [20] Tullis, T. S., & Stetson, J. N. (2004). *A Comparison of Questionnaires for Assessing Website Usability*. In *Usability professional association conference* (pp. 1-12).
- [21] Nielsen, J. (2014). *Demonstrate Thinking Aloud by Showing Users a Video*. Nielsen Norman Group. Version: 01.09.2014
- [22] Keck, M. (2011). *Herausforderungen App-Usability – Auf was Sie bei der Gestaltung von Apps achten sollte*. Version: 17.08.2010
- [23] Ambassadors for Life. Webseite (2017) <http://www.ambassadorsforlife.org/> abgerufen am 28.11.2016
- [24] Google Play. *AZ Screen Recorder*. Webseite (2017) <https://play.google.com/store/apps/details?id=com.hecorat.screenrecorder.free&hl=de>. Version: 10.02.2017
- [25] Frommann, U. (2005). *Die Methode „Lautes Denken“*. Version: 10.08.2015
- [26] Nielsen, J. (2012). *Thinking Aloud: The #1 Usability Tool*. Nielsen Norman Group. Version: 16.01.2012
- [27] Rasch, D., Verdooren, L. R., & Gowers, J. I. (2007). *Planung und Auswertung von Versuchen und Erhebungen*. Oldenbourg Verlag.
- [28] Höst, M., Regnell, B., & Wohlin, C. (2000). *Using Students as Subjects—A Comparative Study of Students and Professionals in Lead-Time Impact Assessment*. *Empirical Software Engineering*, 5(3), 201-214. Springer-Verlag.
- [29] Bangor, A., Kortum, P., & Miller, J. (2009). *Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale*. *Journal of usability studies*, 4(3), 114-123. Usability Professionals` Association

A

Materialien der Benutzerstudie

Code :476183



ulm university universität
uulm

Fakultät Ingenieurwissenschaften, Informatik und Psychologie

Institut für Datenbanken und Informationssysteme (DBIS)

Einverständniserklärung

Die Richtlinien der Deutschen Forschungsgemeinschaft (DFG) sehen vor, dass sich die Teilnehmer/innen an Studien mit ihrer Unterschrift explizit und nachvollziehbar mit den unten genannten Bedingungen einverstanden erklären, um damit zu dokumentieren, dass sie freiwillig an unserer Studie teilnehmen. Aus diesem Grund möchten wir Sie bitten, die vorliegende Einverständniserklärung zu unterzeichnen, bevor Sie an unserer Studie teilnehmen.

Projektleitung: Johannes Schobel (Wissenschaftl. Mitarbeiter)
Universität Ulm – Fakultät für Ingenieurwissenschaften und Informatik
Institut für Datenbanken und Informationssysteme (DBIS)
E-mail : johannes.schobel@uni-ulm.de

Ablauf der Studie und Dauer:

In der vorliegenden Studie werden Sie eine oder mehrere der folgenden Tätigkeiten durchführen: Fragen beantworten, Informationen aus dem Gedächtnis abrufen, Urteile abgeben und bestimmte Aufgabenstellungen an einem mobilen Endgerät ausführen.

Sie nehmen an einer Benutzerstudie teil, wo mehrere verschiedene Teilnehmer die Benutzbarkeit der mobilen Anwendung testen. Während der Studie wird Ihre Stimme und die Oberfläche des mobilen Endgeräts aufgenommen. Es werden keine personenbezogene Daten gespeichert.

Es ist für unsere Studie wichtig, dass Sie die einzelnen Aufgaben und Tätigkeiten in der vorgegebenen Reihenfolge bearbeiten. Die Dauer der Studie liegt bei ca. 30 Minuten.

Vertraulichkeit: Ihre Daten sind selbstverständlich vertraulich und werden nur in anonymisierter Form genutzt. Zu keinem Zeitpunkt der Studie werden wir Sie bitten, Ihren Name oder andere eindeutige Informationen zu nennen. Demographische Angaben, wie Alter oder Geschlecht lassen keinen eindeutigen Schluss auf Ihre Person zu.

Freiwilligkeit: Ihre Teilnahme an dieser Studie ist freiwillig. Es steht Ihnen zu jedem Zeitpunkt frei, Ihre Teilnahme abubrechen, ohne dass Ihnen daraus Nachteile entstehen.

Offene Fragen: Falls Sie noch Fragen zu dieser Studie haben sollten, wenden Sie sich bitte an die Leitung dieser Studie.

Diese Ausfertigung der Einverständniserklärung verbleibt zu Dokumentationszwecken bei der Projektleitung. Sollten Sie eine Kopie für Ihre eigenen Unterlagen benötigen, informieren Sie bitte die Projektleitung darüber.

Einverständniserklärung: Hiermit bestätige ich, dass ich älter als 18 Jahre bin, sowie sämtliche obigen Angaben gelesen und vollständig verstanden habe.

Datum

Unterschrift

Code :476183



ulm university universität
uulm

Fakultät Ingenieurwissenschaften, Informatik und Psychologie

Institut für Datenbanken und Informationssysteme (DBIS)

Demographie

Bitte beantworten Sie zunächst die folgenden Fragen zu Ihrer Person. Die Angaben werden selbstverständlich vertraulich behandelt und dienen lediglich zu Auswertungszwecken!

1. Geschlecht: weiblich männlich

2. Alter: _____ Jahre

3. Welche berufliche Ausbildung trifft am ehesten auf Sie zu?

- Auszubildende[r]/Student[in]
- abgeschlossene Berufsausbildung (Lehre)
- abgeschlossene Ausbildung an einer Meister- oder Technikerschule
- Akademiker[in]
- Sonstiges, und zwar _____

4. Wie hoch ist Ihre aktuelle Anzahl an Ausbildungsjahren (inkl. Grundschule):

5. Welchen höchsten Bildungsabschluss haben Sie?

- ohne Abschluss
- Hauptschulabschluss oder Volksschulabschluss
- Realschulabschluss (Mittlere Reife)
- Fachhochschulreife
- Hochschulreife (Abitur)
- Fachhochschulabschluss
- Hochschulabschluss
- Sonstiger Abschluss, und zwar: _____

6. Besitzen Sie ein Tablet ? ja nein

7. Besitzen Sie ein Smartphone? ja nein

8. Falls ja, welches Betriebssystem hat ihr Smartphone ?

iOS

Android

Windows

Andere

9. Wie hoch würden Sie Ihre Kenntnisse im Bereich Technik selbst einschätzen?

sehr gering

gering

eher gering

weder gering
noch hoch

eher hoch

hoch

sehr hoch

10. Wie regelmäßig nutzen Sie ihr mobiles Endgerät?

sehr gering

gering

eher gering

weder gering
noch oft

eher öfters

oft

sehr oft

11. Wie hoch würden Sie Ihre Kenntnisse im Umgang mit mobilen Applikationen selbst einschätzen?

sehr gering

gering

eher gering

weder gering
noch hoch

eher hoch

hoch

sehr hoch

12. Haben Sie schon Data Collection Apps benutzt / damit zu tun gehabt ?

ja

nein

Code : 476183

Vergleichsfragebogen

1. Welche Darstellung des Fragebogens bevorzugen Sie auf dem mobilen Endgerät, wenn Sie eine Auswahl von Antworten zur Verfügung gestellt bekommen?
 - Ankreuzen der Antwort
 - Grafische Darstellung der Antwortmöglichkeiten (Beispiele aus dem Test)
 - Andere

2. Welche Darstellung bevorzugen Sie auf dem mobilen Endgerät für Fragen mit einer möglichen Antwort?
 - Antwortmöglichkeit ankreuzen
 - Grafische Darstellung der Antwortmöglichkeiten (Beispiele aus dem Test)

3. Welche Darstellung bevorzugen Sie auf dem mobilen Endgerät für Fragen mit mehreren Antwortmöglichkeiten?
 - Antwortmöglichkeiten ankreuzen
 - Grafische Darstellung Ihrer Antwort an einem Slider

4. Welche Darstellung des Fragebogens auf dem mobilen Endgerät war für Sie einfacher zu bedienen?
 - Fragebogen mit Antwortmöglichkeiten zum ankreuzen
 - Fragebogen mit verschiedenen grafischen Darstellungen (Beispiele aus dem Test)

5. Bevorzugen Sie ausfüllen von Fragebogen :
 - Auf Papier
 - Auf mobilen Endgeräten

CODE : 476183

SYSTEM USABILITY SCALE

AUSSAGE	STIMME GAR NICHT ZU				STIMME VOLL ZU
1. Ich kann mir sehr gut vorstellen, das System regelmäßig zu nutzen.					
2. Ich empfinde das System als unnötig komplex.					
3. Ich empfinde das System als einfach zu nutzen.					
4. Ich denke, dass ich technischen Support brauchen würde, um das System zu nutzen.					
5. Ich finde, dass die verschiedenen Funktionen des Systems gut integriert sind.					
6. Ich finde, dass es im System zu viele Inkonsistenzen gibt.					
7. Ich kann mir vorstellen, dass die meisten Leute das System schnell zu beherrschen lernen.					
8. Ich empfinde die Bedienung als sehr umständlich.					
9. Ich habe mich bei der Nutzung des Systems sehr sicher gefühlt.					
10. Ich musste eine Menge Dinge lernen, bevor ich mit dem System arbeiten konnte.					

Abbildungsverzeichnis

Abbildung 1.1: Qualitätsmodell für externe und interne Qualität nach ISO 25010 [3]	2
Abbildung 4.1: Ablauf der Benutzerstudie	23
Abbildung 4.2: Darstellung von der Slidebar unter Frage 5 (links) und die Darstellung der Jahre als Antwortmöglichkeiten (rechts)	25
Abbildung 5.1: Betriebssystem auf dem Smartphone der Teilnehmer.....	30
Abbildung 5.2: Mögliche Lösung für das Problem beim Range-Element	34
Abbildung 5.3: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 1 des Vergleichsfragebogens	35
Abbildung 5.4: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 2 des Vergleichsfragebogens	36
Abbildung 5.5: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 3 des Vergleichsfragebogens	37
Abbildung 5.6: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 4 des Vergleichsfragebogens	38
Abbildung 5.7: Anzahl der Teilnehmer mit der entsprechenden Antwort zur Frage 5 des Vergleichsfragebogens	39
Abbildung 5.8: Bangor Adjective Scale [29]	40
Abbildung 5.9: Ergebnisse des SUS-Scores der einzelnen Teilnehmer.....	40

Tabellenverzeichnis

Tabelle 1: Aufteilung der Versionen der Fragebögen auf die Teilnehmer	29
Tabelle 2: Wichtige Punkte die beim Laut Denken-Protokoll angemerkt wurden.....	32

Name: Jasmina Bandic

Matrikelnummer: 784455

Erklärung

Ich erkläre, dass ich die Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ulm, den.....

Jasmina Bandic