

**Die Wirkung von UX-Gesetzen auf die
Usability und User Experience
interaktiver Webschnittstellen**
Eine empirische Untersuchung

Bachelor Thesis
Naturwissenschaftliche Fakultät, Universität Bern

verfasst von

Romano Brentani

Bern, Schweiz

Betreuung:

PD Dr. Kaspar Riesen
Institut für Informatik (INF)
Universität Bern, Schweiz

Zusammenfassung

Heutzutage interagieren Menschen täglich mit unterschiedlichsten Produkten über eine Schnittstelle, um alltägliche Probleme oder Aufgaben zu lösen. Die Gestaltung von Produkten, die leicht und intuitiv bedienbar sind, also eine gute Usability aufweisen, wird daher bedeutungsvoller. Die Berücksichtigung von Gesetzmässigkeiten in menschlichen Wahrnehmungs- und Informationsverarbeitungsprozessen in Form von Usability-Gesetzen (UX-Gesetzen) gewinnt aus diesem Grund zunehmend an Relevanz. UX-Gesetze wie das Gesetz von Miller, Teslers Gesetz oder das Gesetz von Jakob sind in der Theorie ausführlich dokumentiert. Um ihre tatsächlichen Auswirkungen auf die Usability und User Experience besser zu verstehen, sind jedoch auch empirische Untersuchungen in Form von Usability- und User Experience-Tests mit realen Teilnehmenden notwendig. Durch solche Studien lässt sich untersuchen, wie sich diese Gestaltungsprinzipien auf das Interaktionsverhalten, die kognitive Belastung und die Effizienz der Studienteilnehmenden auswirken und wie sich darauf basierend die Gestaltung von Schnittstellen verbessern und erweitern lässt. Die Arbeit untersucht anhand eines experimentellen Usability- und User Experience-Tests mit 37 Teilnehmenden, wie sich ausgewählte UX-Gesetze auf das Interaktionsverhalten sowie die wahrgenommene Gebrauchstauglichkeit (Usability) einer Webschnittstelle auswirken und welche Rolle dabei mentale Modelle von Personen spielen, welche die Schnittstellen nutzen. Für den Test wurde eine voll funktionsfähige Webseite entwickelt, über welche die Teilnehmenden individuell eine vollständige User Experience durchlaufen. Die Ergebnisse bestätigen, dass UX-Gesetze signifikante Auswirkungen auf das Interaktionsverhalten haben: Aufgaben, die eher auf Erinnern basieren, erfordern zwar mehr Zeit und kognitive Ressourcen als solche, die Erkennung ermöglichen. Dennoch können suchbasierte Interaktionen, die stärker auf Erinnern beruhen, je nach Problemkontext zu einer effizienten Zielerreichung führen. Textbasiertes Feedback mit zu vielen Informationen wird im Sinne des „Paradoxons des aktiven Nutzers“ ignoriert oder ausgeblendet, wenn es die Kapazität des Kurzzeitgedächtnisses überfordert. Die Übertragung von Entscheidungen auf das Produkt oder System reduziert nach dem Gesetz von Tesler die Komplexität und steigert die Effizienz – vorausgesetzt, diese Entscheidungen sind für die Zielgruppe nicht kritisch. Der gezielte Einsatz des Restorff-Effekts (z.B. visuelle Hervorhebungen) kann die Aufmerksamkeit effektiv auf zielführende Elemente lenken und so die Usability steigern. Die Anpassung von Produktdesigns an die mentalen Modelle der Produktnutzenden hilft schliesslich, längere Lernprozesse zu vermeiden, Verständnislücken zu schliessen und die Effizienz zu steigern.

Danksagung

Ich möchte mich herzlich bei meinem Betreuer, PD Dr. Kaspar Riesen, für seine wertvolle Anleitung, konstruktiven Ideen sowie für die fachliche Unterstützung während des gesamten Entstehungsprozesses dieser Arbeit bedanken. Mein besonderer Dank gilt auch allen Teilnehmenden der Studie, die sich die Zeit genommen haben, an der Untersuchung teilzunehmen. Durch ihre engagierte Teilnahme konnten relevante Daten erhoben werden, die wesentlich zur Durchführung und Auswertung dieser Arbeit beigetragen haben.

Inhaltsverzeichnis

1	Einführung	1
1.1	Kontext	1
1.2	Human-Computer Interaction (Begriffe und Definitionen)	2
1.3	Motivation	4
1.4	Forschungsfragen und Ziel der Arbeit	5
1.5	Ausblick	7
2	Gestaltungsgesetze, menschliche Wahrnehmung und Verhaltensmuster	9
2.1	Menschliche Faktoren	9
2.2	Menschliche Verhaltensmuster	12
2.3	Gestaltungsgesetze und Designprinzipien	13
2.3.1	UX-Gesetze	13
2.3.2	Usability-Prinzipien / UX-Prinzipien	18
2.3.3	Weitere Gestaltungselemente	22
3	Messen der Usability und User Experience	24
3.1	Usability-Tests	24
3.2	Usability und User Experience-Metriken	26
3.2.1	Performance-Metriken	26
3.2.2	Self-reported-Metriken	30
3.2.3	Problem-basierte Metriken (Usability-Issues)	32
3.2.4	Kombinierte und vergleichende Metriken	34
4	Durchführung und Evaluation des Usability- und User Experience-Tests	35
4.1	Durchführungsmethode des Usability- und User Experience-Tests	36
4.2	Aufbau und Abfolge des Usability-Tests	37
4.2.1	Technische Implementierung der Testwebseite	39
4.2.2	Aufbau und Funktionsweise der Testwebseite	40

4.3	Evaluierung der Navigationseffizienz	43
4.3.1	Usability und User Experience-Metriken	44
4.3.2	Ergebnisse	44
4.3.3	Interpretation der Ergebnisse	46
4.4	Millers Gesetz	49
4.4.1	Testaufbau	50
4.4.2	Usability und User Experience-Metriken	52
4.4.3	Ergebnisse	52
4.4.4	Interpretation der Ergebnisse	55
4.5	Gesetz von Hick	58
4.5.1	Testaufbau	58
4.5.2	Usability und User Experience-Metriken	59
4.5.3	Ergebnisse	60
4.5.4	Interpretation der Ergebnisse	63
4.6	Das Paradox des aktiven Nutzers	65
4.6.1	Testaufbau	65
4.6.2	Usability- und User Experience-Metriken	68
4.6.3	Ergebnisse	68
4.6.4	Interpretation der Ergebnisse	72
4.7	Gesetz von Tesler	74
4.7.1	Testaufbau	74
4.7.2	Usability und User Experience-Metriken	75
4.7.3	Ergebnisse	75
4.7.4	Interpretation der Ergebnisse	78
4.8	Kontrollierbarkeit und Steuerbarkeit	78
4.8.1	Testaufbau	79
4.8.2	Usability und User Experience-Metriken	80
4.8.3	Ergebnisse	81
4.8.4	Interpretation der Ergebnisse	87
4.9	Gesetz von Jakob und Mentale Modelle	90
4.9.1	Testaufbau	90
4.9.2	Usability- und User Experience-Metriken	93
4.9.3	Ergebnisse	95
4.9.4	Interpretation der Ergebnisse	102
4.10	Restorff-Effekt	110
4.10.1	Testaufbau	110
4.10.2	Usability und User Experience-Metriken	112
4.10.3	Ergebnisse	112

4.10.4	Interpretation der Ergebnisse	113
5	Schlussfolgerungen und künftige Arbeiten	115
5.1	Schlussfolgerungen	115
5.2	Künftige Arbeiten	117
A	Illustrationen zu UX-Gesetzen und Gestaltungsprinzipien	119
A.1	Gesetz von Miller	119
A.2	Gesetz von Jakob	119
A.3	Mentale Modelle	119
A.4	Restorff-Effekt	120
A.5	Steuerbarkeit und Kontrollierbar	121
A.6	Flexibilität und mentale Modelle	121
A.7	Angebotscharakter	121
B	Diagramme der Implementierung	123
B.1	Schichten der Webseite	123
B.2	Komponnetendiagramm Frontend	123
C	Webseite - Screenshots	125
C.1	Anmeldeformular	125
C.2	Umfrageformular	126
C.3	Navigationsstrukturen	126
C.3.1	Linke Seitennavigation (Off-Canvas-Menü)	126
C.3.2	Horizontale Navigation	127
C.3.3	Meg-Drop-Down Navigation	127
C.4	Hauptseite der Webseite	127
C.4.1	Experimente	128
C.4.2	Fortschrittsanzeige	128
C.5	Einstellungen	129
C.6	Versuchsliste	130
C.7	Versuchsdetails	131
D	Versuchsexperimente - Screenshots	132
D.1	Versuchsinstruktionen	132
D.2	Millers Gesetz	133
D.2.1	Versuchsvarianten zum Prinzip Erkennen	133
D.2.2	Suchfunktionen	133
D.3	Gesetz von Hick	134
D.4	Paradox des aktiven Nutzers	137

D.4.1	Dialog, der jederzeit geschlossen werden kann	137
D.4.2	Öffnen des Dialoges	137
D.4.3	Dialog der Gruppe B	137
D.4.4	Tipp-Dialog	137
D.5	Teslers Gesetz	140
D.5.1	Progressive Offenlegung	140
D.6	Kontrollierbarkeit und Steuerbarkeit	141
D.7	Jakob's Gesetz und mentale Modelle	144
D.7.1	Nutzerkonfigurationen	144
D.7.2	Usability-Issue	144
D.8	5 Schnittstellen zur Untersuchung mentaler Modell	147
D.8.1	Off-Canvas-Menü Schnittstellen	147
D.8.2	Mega-Drop-Down Menüschnittstelle	148
D.8.3	Schnittstellen mit textbasierten Menükomponenten (Kein Off-Canvas)	150
D.9	Restorff-Effekt	151
D.9.1	Erster Versuchstest	151
D.9.2	Zweiter Versuchstest	151
D.9.3	Dritter Versuchstest	152

Bibliography **153**

Kapitel 1

Einführung

1.1 Kontext

Die vorliegende Thesis liegt im Forschungsbereich der Human-Computer Interaction (HCI) und verfolgt das Ziel, die Interaktion von Nutzern mit einer Webschnittstelle (Webseite) zu analysieren. Der Forschungsbereich der HCI richtet seinen Fokus primär auf das Zusammenspiel und die Beziehungen zwischen Menschen und Produkten oder Systemen, die über ein Interface bedient werden können und vereint daher Aspekte aus der Psychologie, Informatik, Marketing sowie Neurowissenschaften. Abbildung 1.1 zeigt, aus welchen Bereichen und Disziplinen die HCI vorwiegend ihre Erkenntnisse bezieht. Dabei stehen weniger die physischen Eigenschaften der Systeme oder Geräte im Vordergrund, sondern mehr die kognitiven und mentalen Prozesse sowie Fähigkeiten der Nutzer und die Art und Weise, wie Menschen mit einem Computer interagieren. Ein zentrales Anliegen – sowohl in der HCI allgemein als auch in dieser Arbeit – ist die Berücksichtigung kognitiver Leistungsfähigkeit sowie möglicher Einschränkungen und Beeinträchtigungen der Nutzer, die das Nutzungserlebnis und die Interaktion beeinflussen können. Ein wichtiger Gegenstand der Arbeit ist es, nicht nur einzelne, isolierte Interaktionen mit einem bestimmten Ziel zu betrachten, sondern einen gesamten Interaktionsprozess mit mehreren Handlungszyklen und die damit verbundene Nutzererfahrung einzubeziehen. Um das gesamte Nutzererlebnis betrachten und untersuchen zu können, müssen also auch weitere Aspekte, wie die Erwartungen, Ziele, Einstellungen, Verhaltensweisen und Reaktionen der Nutzer vor, während und nach der Nutzerinteraktion berücksichtigt werden.



Abbildung 1.1: Wichtige Disziplinen und Bereiche, aus denen die HCI ihre Erkenntnisse bezieht

1.2 Human-Computer Interaction (Begriffe und Definitionen)

Aufgrund der zentralen Bedeutung der HCI für diese Arbeit werden im Folgenden zunächst der Begriff der Human-Computer Interaction als Forschungsgebiet beschrieben und anschliessend definiert sowie weitere zentrale und relevante Konzepte und Begriffe vorgestellt, die im Verlauf der Arbeit eine wesentliche Rolle spielen.

Human-Computer Interaction

Die HCI beschäftigt sich mit der Gestaltung von Produkten, Systemen oder Services und mit der Frage, wie Menschen über Schnittstellen mit Produkten oder Systemen interagieren. Als wichtigster Faktor steht der Mensch mit seinen Fähigkeiten, Einschränkungen und Bedürfnissen im Zentrum. Diese menschlichen Aspekte werden in der HCI gezielt berücksichtigt, um Gestaltungsansätze oder Gestaltungsprinzipien zu entwickeln, die sich an Menschen und deren Fähigkeiten orientieren. Ein zentrales Ziel der HCI ist es zu untersuchen, wie Produkte oder Systeme für den menschlichen Gebrauch gestaltet sind und wie diese gestaltet sein sollten, um die Nutzung soweit wie möglich zu vereinfachen. Dabei geht es insbesondere um die Frage, wie gebrauchstauglich ein Produkt ist, also wie leicht es sich erlernen und nutzen lässt. Zudem wird teils auch analysiert, wie effektiv und effizient ein Produkt zur Lösung konkreter Aufgaben eingesetzt werden kann. Heute reicht der Untersuchungsbereich der HCI jedoch über rein funktionale Aspekte hinaus. Im Fokus steht

zunehmend auch die Gestaltung und Entwicklung von Produkten oder Systemen, die nicht nur intuitiv und einfach bedienbar sind, sondern den Nutzern vor, während und nach der Nutzung positive und zufriedenstellende Erfahrungen und Erlebnisse ermöglichen.

Definition HCI

Es gibt unterschiedliche Definitionen des Forschungsfeldes der HCI. Eine für diese Arbeit passende, gebräuchliche und prägnante Definition nach [1] lautet wie folgt: Das Forschungsgebiet der Human-Computer Interaction beschäftigt sich mit der Erforschung, Gestaltung und Bewertung der Interaktion von Mensch und Computer (Hardware und Software).

Usability, User Experience und Useful

Ein zentraler Bestandteil dieser Arbeit ist es zu untersuchen, inwiefern Gestaltungsgesetze (UX-Gesetze) sowie Designprinzipien und -richtlinien aus dem Bereich der HCI die Benutzerfreundlichkeit, auch als Gebrauchstauglichkeit oder Usability bezeichnet, sowie die gesamte Nutzererfahrung mit einem Produkt (User Experience) beeinflussen. Hierfür werden die beiden Begriffe der Usability und User Experience näher erläutert und mit dem Begriff der Usefulness verglichen.

Der Begriff der User Experience fokussiert sich auf den gesamten Interaktionsprozess eines Nutzers mit einem System oder Produkt. Bei der User Experience werden alle Aspekte der Interaktion des Endbenutzers mit einem System betrachtet, unter anderem die Benutzeroberfläche, Grafik, das Produktdesign, wie auch die physische Interaktion. Zudem spielen alle Wahrnehmungen, Erfahrungen, Erwartungen, Reaktionen und Emotionen, die vor, während und nach der Nutzung eines Produktes entstehen, eine bedeutende Rolle in der User Experience und werden für dessen Bewertung berücksichtigt [1].

Bei der Untersuchung der Usability eines Produktes oder Services geht es in der Regel um die Bewertung, wie angenehm, einfach, intuitiv, effizient und effektiv es ist, mit dem Produkt oder System Aufgaben zu erledigen oder Ziele zu erreichen. Usability ist demnach ein Mass dafür, wie gebrauchstauglich ein Produkt oder System ist und bezieht sich unmittelbar auf das Produkt selbst sowie dessen Nutzung. Die User Experience hingegen berücksichtigt zusätzlich Faktoren, die über die reine Produktnutzung hinausgehen und ausserhalb des eigentlichen Produkts liegen können, wie zum Beispiel die Emotionen oder Reaktionen, welche während einer Interaktion mit einem Produkt entstehen können [2]. Abbildung 1.2 veranschaulicht den Unterschied sowie die Beziehung zwischen den Begriffen Usability und

User Experience.

Bei der Usefulness geht es im Gegensatz zur Usability oder User Experience um die Frage, ob mit dem Produkt oder System überhaupt Aufgaben und Probleme gelöst und Ziele erreicht werden können. Das heisst, es geht um die Frage, ob ein Produkt für einen bestimmten Zweck oder für ein bestimmtes Problem tatsächlich verwendet werden kann. An dieser Stelle soll angedeutet werden, dass die Untersuchung der Usefulness von Produkten nicht der Gegenstand der Arbeit ist.

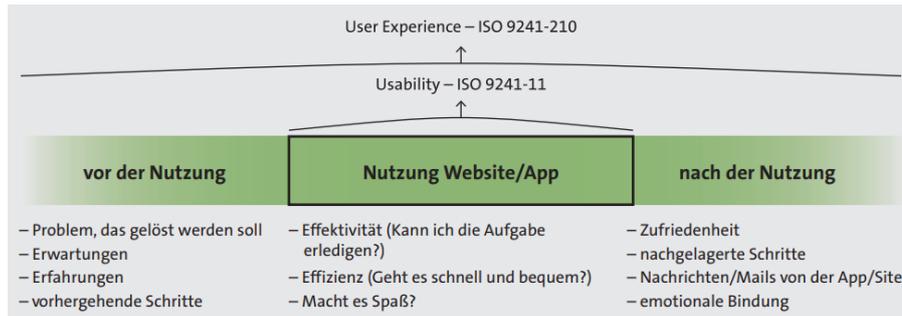


Abbildung 1.2: Usability und User Experience

Nutzende/Benutzer

Der Begriff Nutzer oder Benutzer wurde weiter oben bereits mehrfach verwendet. Dieser Begriff kann abhängig vom Kontext unterschiedliche Bedeutungen haben. In dieser Thesis bezieht sich der Begriff Nutzer oder Benutzer auf die Person, welche ein System, ein Produkt oder Dienstleistung selbst verwendet oder direkt damit in Kontakt steht. Der Begriff des Nutzers oder des Benutzers lässt sich für diese Arbeit entsprechend wie folgt analog definieren: Ein Nutzer oder Benutzer ist eine Person, welche mit einem System, einem Produkt oder Dienstleistungen interagiert oder damit in Kontakt steht [1].

1.3 Motivation

Das Ziel, Schnittstellen von Produkten oder Systemen leicht erlernbar und einfach bedienbar zu gestalten, besteht oft nicht nur bei Web- oder mobilen Anwendungen. Grundsätzlich betrifft dies alle Arten von Produkten, Services oder Dienstleistungen, die über eine Benutzeroberfläche gesteuert werden können. Ein zentrales Anliegen besteht darin, die Benutzeroberfläche oder das Design so verständlich und einfach wie möglich zu gestalten. Dafür gibt es mehrere Gründe: Intuitive und leicht verständliche Produkt- oder Systemdesigns tragen dazu bei, Bedienfehler zu vermeiden und den Lernaufwand zu minimieren. Gleichzeitig wird auch die Zeit reduziert,

die Nutzer benötigen, um Aufgaben erfolgreich zu bewältigen. Die Möglichkeit, effizient und zielgerichtet zu interagieren, ohne vorerst viel Zeit in das Verstehen des Systems investieren zu müssen, ist entscheidend für eine hohe Usability und eine insgesamt zufriedenstellende User Experience.

Diese Eigenschaften sind besonders relevant für Produkte oder Anwendungen, die bei repetitiven Aufgaben zum Einsatz kommen. Nutzer sind häufig nicht bereit, viel Zeit in das Erlernen von Anwendungen zu investieren, die zur Lösung alltäglicher, kleinerer Aufgaben dienen. Diese Tatsache wird im nächsten Kapitel als das sogenannte Paradox des aktiven Nutzers beschrieben. Mögliche Beispiele sind das Versenden von E-Mails oder das Online-Bezahlen von Rechnungen. Aus diesem Grund ist es besonders wichtig, dass solche Produkte intuitiv bedienbar sind und sich mit möglichst geringem Lernaufwand nutzen lassen. Würde der Lernaufwand solcher Produkte zu hoch ausfallen, so könnte dies möglicherweise dazu führen, dass viele Nutzer abspringen und ein anderes Produkt verwenden würden, weil sie denken, sie verschwenden während der Nutzung ihre Zeit. Doch um intuitive und einfach erlernbare und verständliche Produkte oder Systeme zu entwickeln, ist es notwendig, ein Verständnis dafür zu haben, wie Menschen Funktionen wahrnehmen, erlernen, nutzen und sich einprägen und welche Erwartungen sie an ein Produkt haben. Auf der anderen Seite ist es aber auch von Bedeutung zu verstehen, wie die Produktgestaltung auf die menschliche Wahrnehmung wirkt und welche Arten von Produktgestaltungen den Menschen in den oben genannten Aspekten unterstützen.

Hier setzen die Gesetze der Usability und User Experience an. Sie liefern Erkenntnisse aus Psychologie, Kognitionswissenschaft und Gestaltungspraxis, die helfen können, menschliches Verhalten und die Wahrnehmung besser zu verstehen und gezielt in die Gestaltung von Schnittstellen einzubeziehen. Durch die Anwendung dieser Prinzipien können Entwickler von Produktoberflächen Designentscheidungen treffen, die nicht nur die Effizienz und Verständlichkeit verbessern, sondern auch die Zufriedenheit der Nutzer langfristig sichern. Sie bieten eine strukturierte Grundlage, um Komplexität zu reduzieren und Interaktionen so zu gestalten, dass Nutzer ihr bereits vorhandenes Wissen für die Nutzung neuer Produkte einsetzen können.

1.4 Forschungsfragen und Ziel der Arbeit

Die Arbeit verfolgt insgesamt zwei Hauptziele. Jedes dieser Ziele behandelt schliesslich mehrere zentrale Kernfragen, die im Folgenden beschrieben werden.

Das erste Ziel dieser Arbeit ist es, Auswirkungen unterschiedlicher Gestaltungsgesetze (UX-Gesetze), auch als Gestaltungsregeln oder Gestaltungsprinzipien bezeichnet, auf die Nutzerinteraktion und das Nutzerverhalten qualitativ bzw. empi-

risch zu untersuchen. Dies erfolgt im Rahmen eines Usability- und User Experience-Tests mit realen Nutzern. Im Mittelpunkt steht die Frage, wie sich verschiedene UX-Gesetze auf die menschliche Wahrnehmung, das (Fehl)-Verhalten sowie auf die Effizienz und kognitive Belastung der Nutzung auswirken. Darüber hinaus sollen während der Durchführung des Usability-Tests konkrete Usability-Probleme (Usability-Issues) identifiziert werden. Dabei wird insbesondere untersucht, welche dieser Probleme auf eine Diskrepanz zwischen dem Verständnis und den Erwartungen der Nutzenden und den Annahmen der Designer zurückzuführen sind. Denn solche Differenzen können ein Hinweis auf Verständniskluffen sein, welche die Nutzung erschweren können. Ein besonderer Fokus liegt auch auf der Evaluation unterschiedlicher Webschnittstellen. Es wird untersucht, inwiefern verschiedene Strukturen und Designs die Erlernbarkeit, Einprägsamkeit, kognitive Belastung und Nutzungseffizienz beeinflussen und inwieweit sie die Übertragung bestehenden Wissens sowie vorhandener Erfahrungen unterstützen. Dabei stellt sich die Frage, ob die Einhaltung gängiger Designkonventionen zu einer effizienteren, fehlerärmeren und insgesamt angenehmeren Interaktion führt als bei Webschnittstellen, die von verbreiteten Gestaltungskonventionen und Gestaltungsnormen abweichen. Gleichzeitig wird untersucht, ob sich durch ein selbstbeschreibendes Design mit starkem Angebots- und Aufforderungscharakter auch unkonventionelle Webschnittstellen intuitiv, leicht erlernbar und nutzbar gestalten lassen. Dabei wird ebenfalls analysiert, ob und inwieweit solche Schnittstellen trotz abweichender Strukturen und Muster ein positives Nutzererlebnis ermöglichen können.

Das zweite Ziel, welches die Arbeit verfolgt, ist die Untersuchung des aktiven Interaktionsverhaltens der Nutzer sowie dessen Veränderungen und Entwicklungen im Verlauf der Nutzung einer voll funktionsfähigen Website. Der erste Untersuchungsaspekt, welcher im Rahmen dieses zweiten Ziels erfolgt, ist der Vergleich von heute teils weit verbreiteten Navigationsstrukturen, welche für die erwähnte Webschnittstelle implementiert wurden und entsprechend von den Teilnehmenden während des gesamten Usability-Tests eingesetzt werden müssen. Es soll ermittelt werden, welche Navigationskonzepte sich rasch erlernen und einprägen lassen und welche den Nutzern eine effiziente Navigation ermöglichen. Dabei steht auch die Frage im Raum, welche Navigationsstrukturen am ehesten den Erwartungen und bisherigen Erfahrungen der Nutzer entsprechen und inwiefern sie es ermöglichen, vorhandenes Wissen aus der Nutzung anderer Anwendungen oder Produkte auf die aktuelle Nutzung zu übertragen. Das zweite Ziel zieht sodann einen weiteren Untersuchungsgegenstand nach sich: Die Beobachtung, wie sich das aktive Verhalten der Nutzer im Laufe ihrer Nutzungserfahrung mit der Webseite entwickelt bzw. verändert. Dabei ist es beabsichtigt, dass die Teilnehmenden ein gewohntes, natürliches Interaktions-

verhalten beibehalten. So wird untersucht, ob durch unterschiedliche, teils gezielte Informationsarchitekturen ein Verhalten hervorgerufen werden kann, welches vom Standardnutzungsverhalten abweicht, zugunsten einer angenehmeren, effizienteren und insgesamt zufriedenstellenderen User Experience.

Die Beantwortung der eben beschriebenen Fragen soll letztlich helfen, ein fundiertes Verständnis darüber zu gewinnen, wie gestalterische Entscheidungen in der Schnittstellengestaltung die Interaktionserfahrung sowohl kurzfristig als auch langfristig beeinflussen.

1.5 Ausblick

Die verbleibende Arbeit gliedert sich wie folgt:

Der zentrale Faktor im Forschungsbereich der HCI ist der Mensch. Denn die Art und Weise, wie ein Mensch mit einem System interagiert, hängt stark davon ab, wie er ein Produkt oder ein Produktdesign wahrnimmt und die bereitgestellten Informationen verarbeitet. Aus diesem Grund werden im ersten Abschnitt von Kapitel zwei wichtige Prozesse der menschlichen Informationsverarbeitung- und Wahrnehmung, welche für die spätere Evaluierung des Usability- und User Experience-Tests von Bedeutung sind, kurz erläutert.

Anschliessend werden bekannte Gesetzmässigkeiten in der menschlichen Wahrnehmung in Form von Gestaltungsregeln, sogenannten UX-Gesetzen, vorgestellt. Es wird beschrieben, wie diese Regeln aus gestalterischer Sicht zur Verbesserung und Erhaltung der Usability und somit auch zur Verbesserung der gesamten User Experience beitragen können. Der letzte Abschnitt von Kapitel zwei widmet sich weiteren Gestaltungsrichtlinien. Diese sind weniger als feste Regeln, sondern vielmehr als Prinzipien zu verstehen. Diese können Produktnutzern helfen, die Funktionsweise eines Systems schneller zu erfassen und zu verstehen.

Das dritte Kapitel behandelt die Messung und Beurteilung von Usability und User Experience. Zunächst werden gängige Methoden zur Durchführung entsprechender Studien zur Usability und User Experience mit Teilnehmern vorgestellt. Daraufhin folgt eine Übersicht über mögliche Metriken, mit deren Hilfe sich die Usability und User Experience messen, erfassen und bewerten lassen.

Das vierte Kapitel beschreibt zunächst die Methode des durchgeführten Usability- und User Experience-Tests. Anschliessend wird erläutert, wie der Test konkret abläuft und durchgeführt wird. Danach folgt eine Beschreibung der technischen Implementierungen, die für den Usability- und User Experience-Test vorgenommen wurden. Es folgt die Beschreibung der praktischen Umsetzung der in Kapitel zwei erläuterten Konzepte und Theorien in Form von Experimenten. Dabei wird auch

aufzeigt, welche Metriken während der Experimente eingesetzt wurden, um die Usability und User Experience zu bewerten. Abschliessend werden im vierten Kapitel die Ergebnisse des Tests dargestellt und illustriert.

Kapitel 2

Gestaltungsgesetze, menschliche Wahrnehmung und Verhaltensmuster

In diesem Kapitel geht es in einem ersten Schritt um den Faktor Mensch. Er ist wie bereits erwähnt, ein sehr wichtiger Faktor im Bereich der HCI. Aus diesem Grund werden auf das Interaktionsverhalten teils stark einflussnehmende Wahrnehmungs- und Informationsverarbeitungsprozesse von Nutzern während der Interaktion mit einem System oder Produkt beschrieben. Zusätzlich wird auf ein Benutzerverhalten eingegangen, welches während der Interaktion mit Produkten oder Systemen häufig in Erscheinung tritt: Das sogenannte Paradox des aktiven Nutzers.

In einem zweiten Schritt gibt dieses Kapitel einen Überblick über verschiedene Gestaltungsprinzipien und -regeln, genannt UX-Gesetze und UX-Prinzipien. Es wird erläutert, welches die jeweiligen Kernideen dieser Regeln sind und inwiefern sie die Usability verbessern oder erhalten und damit die Interaktion eines Menschen, also die User Experience mit einem System oder Produkt insgesamt angenehmer gestalten können.

2.1 Menschliche Faktoren

Die Art und Weise, wie Menschen mit Produkten oder Systemen interagieren, hängt davon ab, wie die Menschen das Design oder die Schnittstelle eines Produktes oder Systems wahrnehmen, sich einprägen und wie die vom Design erzeugten Informationen verarbeitet werden. Abhängig davon, wie Informationen wahrgenommen und verarbeitet werden, werden schliesslich Entscheidungen getroffen, welche die Interaktion beeinflussen oder steuern. Aufgrund dieser Tatsachen ist es für die Gestaltung

von Produktdesigns von Vorteil zu verstehen, wie Menschen ein Design wahrnehmen und wie sich die Gestaltung eines Designs auf die menschliche Wahrnehmung auswirkt. Deshalb werden im nächsten Abschnitt die für im Rahmen dieser Arbeit durchgeführten Interaktionen wichtigsten menschlichen Wahrnehmungs- und Informationsverarbeitungsprozesse kurz beschrieben.

Erinnern und Erkennen

Sowohl das Erinnern als auch das Erkennen sind kognitive Prozesse, die mit den Gedächtnis- und Wahrnehmungsfunktionen des Menschen in Verbindung stehen. Beide Vorgänge sind notwendig, um Informationen aus dem Gedächtnis abzurufen.

Grundsätzlich werden drei Arten des Gedächtnisses unterschieden: das sensorische Gedächtnis, das Kurzzeitgedächtnis und das Langzeitgedächtnis. Das sensorische Gedächtnis lässt sich wiederum in den ikonischen Speicher, der visuelle Reize verarbeitet, und den echoischen Speicher, der auditive Reize speichert, unterteilen. Die Kapazität aller Gedächtnisarten wird in der Regel in sogenannten Chunks angegeben. Das sind Informationseinheiten wie Wörter, Zahlen oder Buchstaben. Reize, die eine ausreichende Intensität aufweisen, können durch Erkennungsprozesse in das Kurzzeitgedächtnis überführt werden, wo sie für eine begrenzte Zeit gespeichert und weiterverarbeitet werden. Die dort gespeicherten Informationen zerfallen in der Regel nach etwa 15 bis 30 Sekunden. Störungen, wie beispielsweise Ablenkungen, können diesen Zerfallsprozess beschleunigen. Das Kurzzeitgedächtnis dient somit der kurzfristigen Speicherung und dem schnellen Zugriff auf Informationen. Im Vergleich dazu verfügt das Langzeitgedächtnis über eine deutlich grössere Speicherkapazität. Nach aktuellem Forschungsstand zerfallen einmal gespeicherte Informationen im Langzeitgedächtnis in der Regel nicht. Allerdings benötigt das Speichern von Informationen im Langzeitgedächtnis mehr Zeit. So können dort durchschnittlich etwa acht Chunks pro Sekunde aufgenommen werden [3].

Unter dem Erinnern wird das Abrufen von Informationen, vorwiegend aus dem Langzeitgedächtnis, verstanden. Beim Erkennen geht es dagegen in der Regel darum, ein Ereignis oder Informationen mithilfe von Hinweisen oder anderen äusseren Reizen (meist visuell oder auditiv) als bekannt oder vertraut wahrzunehmen. Die Informationsabrufung durch das Erinnern erfolgt mit viel weniger Hinweisen oder Reizen, als dies beim Erkennen der Fall ist. Dies führt dazu, dass ein Erinnerungsprozess in der Regel mehr kognitive Anstrengung benötigt als ein Erkennungsprozess. Ein Erkennungsprozess kann ein Erinnerungsprozess auslösen und umgekehrt.

Die beiden Gedächtnisfunktionen Erkennen und Erinnern sind bezüglich der Gestaltung von Mensch-Computer-Interaktionen wichtig, denn sie tragen massgebend

dazu bei, wie kognitiv belastend eine Interaktion mit einem System ausfallen kann. Bei der Gestaltung von Interaktionen sollte stets darauf geachtet werden, dass das Kurzzeitgedächtnis nicht überlastet wird und zu viel Information auf einmal gespeichert oder abgerufen werden muss. Dieses Prinzip wird später als das Gesetz von Miller beschrieben, ein Beispiel eines ersten UX-Gesetzes. Erlaubt das Interaktionssystem den Benutzern, den Funktionsumfang des Systems vorwiegend zu erkennen, so muss weniger Information gespeichert und abgerufen und die kognitive Belastung während der Interaktion kann somit reduziert werden [4].

Gezielte Aufmerksamkeit

Gezielte bzw. selektive Aufmerksamkeit beschreibt die Fähigkeit, irrelevante Ereignisse oder Objekte auszublenden und den Fokus auf relevante Ziele oder Informationen zu richten. Um einem Ereignis Aufmerksamkeit zu widmen, werden kognitive Ressourcen benötigt. Da diese beschränkt sind, ist es wichtig, diese für die relevanten Ereignisse zu nutzen. Die selektive Aufmerksamkeit ist demnach eine entscheidende Fähigkeit, um Aufgaben effizient und effektiv zu bewältigen und um die kognitive Belastung während Prozessen reduzieren zu können.

Eine gezielte Aufmerksamkeitslenkung ermöglicht es Nutzern, Probleme schneller zu erkennen und zu lösen, insbesondere im Umgang mit komplexen Produkten oder Systemen [5]. Ein mögliches Beispiel hierfür ist die Interaktion mit einer Webseite, etwa beim Ausfüllen eines Online-Formulars. Während auf der Website möglicherweise Werbung, Pop-ups oder andere ablenkende Elemente in Erscheinung treten, konzentriert sich die gezielte Aufmerksamkeit auf die relevanten Eingabefelder und Anweisungen oder Informationen. Nutzer, die ihre Aufmerksamkeit erfolgreich steuern, können das Formular effizienter ausfüllen, Fehler vermeiden und schneller zum gewünschten Ergebnis gelangen. Ein weiteres Beispiel zeigt sich bei der Nutzung einer Suchmaschine wie zum Beispiel Google. Nach der Eingabe eines Suchbegriffs erscheint eine möglicherweise lange Liste von Suchergebnissen, teils ergänzt durch Werbeanzeigen, Bilder oder auch weiterführende Links. Um relevante Informationen filtern zu können, ist es entscheidend, dass gesuchte Informationen schnell und ohne unnötige Ressourcenverschwendung oder kognitive Belastung lokalisiert werden können.

Entscheiden

Während der Interaktion mit einem System oder Produkt treffen Nutzer in der Regel viele Entscheidungen. Diese basieren auf individuellen Zielen, Erwartungen, Aufgaben oder Bedürfnissen. Häufig stellen sie eine Reaktion auf Reize dar, et-

wa nach einem Erkennungs- oder Erinnerungsprozess. So könnte ein Nutzer eines Online-Shops, nachdem er die Suchfunktion erkannt hat, entscheiden, diese anstelle eines Navigationsmenüs zu verwenden, um gezielt nach einem Produkt zu suchen.

Bei der Gestaltung von Benutzeroberflächen und Interaktionsprozessen sollte berücksichtigt werden, dass das Erkennen eines Zustands nicht automatisch zu einer angemessenen oder korrekten Entscheidung führt [1]. Zudem erfordern Entscheidungen kognitive Ressourcen und Zeit. Deshalb kann es sinnvoll sein, nur dann Entscheidungen von Nutzern einzufordern, wenn diese tatsächlich notwendig sind. In bestimmten Kontexten, etwa im Bereich von Spielen, kann eine häufige Entscheidungsfindung jedoch gewünscht und von Bedeutung für das Nutzungserlebnis sein. Die Komplexität eines Systems lässt sich gezielt verringern, indem Entscheidungen, sofern sie nicht kritisch oder individuell notwendig sind, vom System übernommen werden. Dies reduziert die kognitive Belastung der Nutzer. Je weniger Entscheidungen ein System erfordert, umso einfacher und angenehmer wird die Nutzung wahrgenommen. Das trägt wesentlich zu einer höheren Usability und einer insgesamt positiveren User Experience bei.

2.2 Menschliche Verhaltensmuster

Die Interaktion eines Menschen mit einem Produkt, System oder Anwendung ist vom menschlichen Verhaltensmuster geprägt. Das Verhalten selbst wiederum ist von Bedürfnissen, Zielen und Erfahrungen der Menschen abhängig.

Das Paradox des aktiven Nutzers

Das Paradoxon des aktiven Nutzers beschreibt die häufige Situation, dass Nutzer ein Produkt verwenden, ohne davor viel Zeit investiert zu haben, um zu verstehen, wie das Produkt tatsächlich funktioniert oder welche Funktionalitäten das System bietet. Benutzer lesen also oft keine Dokumentationen, Hilfestellungen oder ändern Standardeinstellungen. Der Wunsch, unmittelbar mit dem Produkt zu arbeiten und Ziele zu erreichen, überwiegt oftmals den Wunsch, das gesamte Produkt oder System und damit alle möglichen Wege zur Erreichung der Ziele zu erkunden. Das heißt, Produktnutzer sind oft motiviert, unmittelbar mit dem Erledigen ihrer Aufgaben zu beginnen und hierfür auf für sie bereits bekannte Vorgehensweisen zurückzugreifen, auch wenn auf diese Weise das Risiko besteht, auf Hindernisse zu stossen [6].

Das ist ein Paradoxon, denn würde zu Beginn genügend Zeit in das Kennenlernen des Produktes investiert, so würden unterschiedliche Vorgehensweisen oder Möglichkeiten zur Zielerreichung kennengelernt. Lang- oder kurzfristig würde dies die Effizienz der Nutzer erhöhen, da auf diese Weise die möglicherweise effizientesten

Wege der Zielerreichung erkundet werden könnten [4]. Aus diesem Verhalten lässt sich ableiten, dass Produkte oder Dienstleistungen so gestaltet sein sollten, dass ihre Funktionsweise intuitiv erkennbar ist (Selbstbeschreibungsfähigkeit) und eine klare Benutzerführung über den gesamten Verlauf der Nutzung hinweg gewährleistet wird.

2.3 Gestaltungsgesetze und Designprinzipien

Um Designprozesse, Designentscheidungen, die menschliche Wahrnehmung von Design sowie die Wirkung von Design auf den Menschen besser zu verstehen und bewusst steuern zu können, kommen sogenannte Gestaltungsgesetze, oft auch Designregeln oder Designprinzipien genannt, zum Einsatz. In der Praxis spricht man heute meist von UX-Gesetzen anstatt von Designregeln. Diese beschreiben Gesetzmässigkeiten zur menschlichen Wahrnehmung von Designs, die helfen, intuitivere und einfachere Design- oder Gestaltungsentscheidungen zu treffen. Aus praktischer Sicht dienen sie also dazu, die Usability gezielt zu erhalten oder zu erhöhen und damit die User Experience positiv zu beeinflussen [7].

Design

Der Begriff Design im Kontext der HCI bezieht sich auf die Gestaltung von Systemen oder Produkten mit dem Ziel, eine hohe Usability sowie eine überzeugende User Experience, also Nutzererfahrung, zu erzielen [1]. Design ist nicht nur eine gestalterische Tätigkeit im ästhetischen Sinne, sondern umfasst auch funktionale und nutzerzentrierte Aspekte. Insbesondere beschäftigt sich Design auch mit dem Zusammenspiel aller Elemente, welche für die Gestaltung eines Systems eingesetzt werden. Entsprechend werden teils auch die beiden Begriffe Usability-Design oder User Experience-Design (UX-Design) eingesetzt.

2.3.1 UX-Gesetze

Im folgenden Abschnitt werden nun verschiedene UX-Gesetze beschrieben. Diese sind nicht als feste Regeln, wie Produkte gestaltet sein sollten, aufzufassen, sondern mehr als Gesetzmässigkeiten in der menschlichen Wahrnehmungs- und Informationsverarbeitung. Die nachfolgend vorgestellten UX-Gesetze werden detailliert in [4] beschrieben.

Gesetz von Miller

Millers Gesetz bezieht sich auf die begrenzte Menge an Informationen, die Menschen gleichzeitig im Kurzzeitgedächtnis behalten können. Laut dem Kognitionspsychologen George Miller speichert unser Gedächtnis Informationen in sogenannten Einheiten oder Chunks. Die bekannteste Aussage, die mit Millers Gesetz häufig in Verbindung gebracht wird, ist die sogenannte magische Sieben: Der Mensch kann etwa sieben (plus/minus zwei) dieser Chunks gleichzeitig im Kurzzeitgedächtnis behalten [8]. Diese Aussage ist allerdings eine vereinfachte Interpretation und nicht der Kerninhalt von Millers Gesetz.

Miller wollte damit nicht sagen, dass das Kurzzeitgedächtnis exakt sieben Einheiten aufnehmen kann. Vielmehr wollte er darauf hinweisen, dass die mentale Kapazität der Menschen begrenzt ist. Diese Begrenzung lässt sich aber nicht auf eine konstante Zahl beschränken. Im Kern geht es beim Gesetz von Miller um die Menge an mentalen Ressourcen, die Menschen aufwenden müssen, um Informationen zu verarbeiten, also um die sogenannte kognitive Belastung [9]. Wenn während der Interaktion mit einem System oder Produkt zu viele mentale Ressourcen gleichzeitig beansprucht werden, kann dies schnell zu einer Überforderung führen. Als Folge haben Nutzer Mühe, der Interaktion zu folgen, übersehen Details oder scheitern an einfacheren Aufgaben. Ein Beispiel dafür ist das Lesen eines Textes. Wenn ein Text lang, unstrukturiert und ohne erkennbare Hierarchie aufgebaut ist, wird das Lesen schnell anstrengend und kognitiv belastend. Ist der Text dagegen gut formatiert, gegliedert und die Zeilenlänge angemessen, wird die Lesbarkeit verbessert und die mentale Belastung reduziert. Dieses Beispiel ist in Anhang A, Abbildung A.2 illustriert.

Gemäss Millers Gesetz ist es demnach für die Nutzer von Vorteil, Produkte oder Systeme so zu designen, dass die kognitive Belastung (Anzahl benötigter mentaler Ressourcen) bei deren Nutzung so gering wie möglich ausfällt.

Gesetz von Hick

Das Gesetz von Hick besagt, dass die benötigte Zeit, um eine Entscheidung in einer Handlungssequenz zu treffen, abhängig von der Anzahl der Auswahlmöglichkeiten und deren Komplexität steigt oder sinkt. Weiter oben wurde bereits verdeutlicht, dass ein Mensch nach Millers Gesetz einer kognitiven Belastung während jeder Interaktion ausgesetzt ist. Je mehr Auswahlobjekte den Nutzerinnen und Nutzern angeboten werden, desto länger wird auch der Entscheidungsprozess ausfallen, sofern eine Entscheidung vom Nutzer verlangt wird. Diese Tatsache ist eine Konsequenz des Auswahlparadox, welches in [10] ausführlich beschrieben wird. Demzufolge steigt

auch die kognitive Belastung, welche nach dem Gesetz von Miller zu Überforderung führen kann.

Unter dem Gesetz von Miller wurde die magische sieben erwähnt, welche oft als maximale Anzahl an Auswahlobjekten für Navigationen oder Menüs in Anwendungen interpretiert wird. Grundsätzlich ist aber nicht davon auszugehen, dass eine Anwendung, zum Beispiel eine Webanwendung, nur noch 7 (plus/minus zwei) Auswahlobjekte aufweisen darf, damit das Kurzzeitgedächtnis nicht überfordert wird. Denn dies würde das Abstraktionslevel der Komplexität in einigen Fällen zu stark reduzieren, was nach dem Gesetz von Tesler die Funktionalität des Systems beeinträchtigen würde.

Ist jedoch die Reaktionszeit von entscheidender Bedeutung, empfiehlt es sich gemäss dem Gesetz von Hick, die Anzahl der Auswahlmöglichkeiten so gering wie möglich zu halten oder Möglichkeiten zu bieten, Optionen zu reduzieren, zum Beispiel durch Filtern. Eine reduzierte Anzahl an Optionen verkürzt die Entscheidungszeit pro Auswahlobjekt, wodurch die kognitive Belastung sinkt. In der Folge verringert sich die Komplexität des Systems oder Produkts insgesamt, was eine schnellere und effizientere Entscheidungsfindung ermöglicht und die Usability erhöht.

Gesetz von Jakob und mentale Modelle

Das Gesetz von Jakob, oft auch als Jakob's Law of the Internet User Experience bezeichnet, besagt, dass Menschen „Gewohnheitstiere“ sind. Sie bevorzugen es, mit vertrauten Objekten zu interagieren und bekannte Dienstleistungen oder Produkte in Anspruch zu nehmen. Durch Erfahrungen und vorhandenes Wissen entwickeln die Benutzer Erwartungen an Produkt- oder Systemdesigns. Diese Erkenntnis ist für die Gestaltung von Benutzeroberflächen, wie etwa Webschnittstellen oder mobilen Schnittstellen, von zentraler Bedeutung und sollte bei Designentscheidungen entsprechend berücksichtigt werden.

Wenn Designer bei der Gestaltung eines Systems auf bekannte und verbreitete Konventionen zurückgreifen, entsteht bei der Nutzung ein Gefühl der Vertrautheit. Dadurch wird die Reibung, die bei der Nutzung digitaler Dienstleistungen entstehen kann, reduziert. Nutzer können sich leichter auf ihre eigentlichen Aufgaben und Ziele konzentrieren, weil sie intuitiv verstehen, wie das Produkt zu bedienen ist. Dadurch wird ein aufwändiger und möglicherweise anstrengender Lernprozess weitestgehend verhindert. Die mentale Anstrengung, die nötig wäre, um die Schnittstelle zu verstehen, verringert sich, was insgesamt zu einer geringeren kognitiven Belastung führt, die Usability erhöht und insgesamt die User Experience verbessert. Wichtig zu beachten ist jedoch: Das Gesetz von Jakob besagt nicht, dass Reibung vollständig zu vermeiden ist. Denn jeder Prozess enthält einen minimalen Grad an Komplexität.

xität, der unvermeidbar ist. Diese grundlegende Tatsache wird später durch das Gesetz von Tesler beschrieben, welches im nächsten Abschnitt thematisiert wird. Bevor darauf eingegangen wird, soll vorerst noch eine wichtige Konsequenz aus dem Gesetz von Jakob erläutert werden: Sogenannte mentale Modelle.

Der Usability-Experte Jakob Nielsen beschrieb im Jahr 2000, dass Nutzer auf Basis ihres angesammelten Wissens und ihrer bisherigen Erfahrungen Erwartungen an Designs entwickeln — insbesondere daran, wie diese funktionieren sollten [4]. Das führte zum Begriff des mentalen Modells eines Nutzers. Ein mentales Modell repräsentiert, welche Erwartungen eine Person an die Funktionsweise eines Systems hat und wie sie denkt zu wissen, wie dieses System funktioniert oder funktionieren sollte und was sie bereits über das System weiss [11].

Das mentale Modell eines Menschen entsteht und verfeinert sich im Laufe der Zeit durch die wiederholte Interaktion mit denselben oder unterschiedlichen Produkten oder Anwendungen. Es ist anzunehmen, dass Nutzer ihre mentalen Modelle auf neue oder ähnliche Systeme übertragen, wann immer ihnen dies möglich erscheint. Das heisst, Wissen aus früheren Erfahrungen wird gezielt genutzt, um die Reibung, die beim Erlernen neuer Systeme entsteht, so weit wie möglich zu minimieren. Die Existenz mentaler Modelle stellt Designer vor die Herausforderung, die Kluft zwischen ihren eigenen Vorstellungen – dem sogenannten Implementationsmodell – und den Erwartungen der Nutzer – den mentalen Modellen der Nutzer – zu verringern oder zu überbrücken. Daraus ergibt sich die Annahme, dass sich Lernprozesse erleichtern und die Zahl an Irrtümern reduzieren lässt, wenn das Design gezielt an die mentalen Modelle der Nutzer angepasst wird.

Durch eine solche Anpassung können bestehende Wahrnehmungs- und Verarbeitungsmuster gezielt berücksichtigt werden. Auf diese Weise wird es den Nutzenden ermöglicht, vorhandenes Wissen sowie frühere Nutzungserfahrungen auf neue Produkte zu übertragen, was den Lernprozess sowie die Einprägsamkeit der Bedienung erleichtert und beschleunigt. Eine Illustration einer Schnittstelle, welche das mentale Modell von Nutzern von Webschnittstellen möglicherweise durch Designkonventionen berücksichtigt, wird in Anhang A, Abbildung A.3 gezeigt.

Gesetz von Tesler und progressive Offenlegung

Teslers Gesetz beschäftigt sich mit dem Prinzip der unvermeidbaren Komplexität in Systemen und Interaktionsprozessen. Es besagt, dass jeder Interaktionsprozess einen gewissen Grad an Komplexität aufweist, der weder vollständig eliminiert noch verhindert werden kann. Diese Komplexität muss entweder vom System selbst oder von den Nutzern bewältigt werden.

Die Berücksichtigung von Teslers Gesetz ist für die Gestaltung von Mensch-

Computer-Interaktionen von zentraler Bedeutung. Es ermöglicht eine Einschätzung darüber, welche Komplexität für einen bestimmten Interaktionsprozess tatsächlich notwendig ist und in welcher Form diese reduziert oder abstrahiert werden kann, ohne die Funktionalität des Gesamtproduktes zu beeinträchtigen. Teslers Gesetz leistet einen wichtigen Beitrag zur Gestaltung benutzerfreundlicher Systeme. Es hilft Designern von Produkten dabei, unnötige Komplexität zu erkennen und zu reduzieren. Dies ist ein entscheidender Prozess, um die kognitive Belastung der Nutzer zu verringern. Dadurch wird die Interaktion nicht nur effizienter, sondern auch angenehmer und besser an die tatsächlichen Bedürfnisse der Zielgruppe angepasst.

Eine Anwendung des Gesetzes von Tesler ist die sogenannte progressive Offenlegung, oft auch Progressive Disclosure (PD) genannt. Bei diesem Prinzip des Interaktionsdesigns werden standardmässig nur relevante Informationen, Aktionen oder Inhalte visualisiert. Zusätzliche Informationen oder Funktionen sind jedoch leicht erkennbar, abrufbar oder auffindbar oder werden schrittweise eingeführt. Damit wird der Komplexitätsgrad nicht direkt verringert, jedoch verschoben. Dies hilft, die Schnittstelle zeitweise schlanker zu visualisieren und entlastet somit den Benutzer bei Interaktionsschritten. Für die progressive Offenlegung gibt es zahlreiche Beispiele. Ein einfaches Beispiel ist ein klassisches Drop-Down-Menü, wie es in Webanwendungen oft eingesetzt wird. Diese Art von Menü versteckt gewöhnlich alle Aktionselemente vorerst. Bei einem Klick auf das Drop-Down-Menü werden die Aktionselemente dann visualisiert. Bei einem erneuten Klick auf das Menü werden die Elemente wieder versteckt. Auf diese Weise wird ermöglicht, dass der Nutzer selbst entscheiden kann, wann er die Informationen sehen will, um eine entsprechende Aktion auszulösen. Ein weiteres Beispiel für eine Komponente, die dieses Prinzip umsetzt, ist das Off-Canvas-Menü [2]. Ähnlich wie bei einem Drop-Down-Menü sind auch hier die Aktionselemente im Ausgangszustand verborgen. Erst durch einen Klick auf das Menü-Icon wird der Menüinhalt eingeblendet, typischerweise durch ein seitliches Hereinschieben von rechts, wodurch die Optionen sichtbar werden. Ein Beispiel einer Off-Canvas-Menüschnittstelle ist in Anhang C unter D.8.1 illustriert.

Restorff Effekt

Ursprünglich besagte der Restorff-Effekt im Zusammenhang mit Designgestaltung, dass von mehreren ähnlich ausschauenden Objekten oder Elementen dasjenige am ehesten im Gedächtnis bleibt, das sich deutlich von den anderen unterscheidet. Heute beeinflusst die Berücksichtigung dieses Effekts weitaus mehr Aspekte des Systemdesigns. Unter dem Restorff-Effekt wird nicht mehr nur die Unterscheidbarkeit einzelner Objekte verstanden, sondern das Zusammenspiel von menschlichem Seh-

vermögen, kognitiver Verarbeitung und Wahrnehmung wird ebenfalls berücksichtigt [12].

Es ist mittlerweile gut erforscht, dass Menschen in der Lage sind, Unterschiede zwischen Objekten innerhalb von Sekundenbruchteilen zu erkennen. Darüber hinaus verfügen wir über die Fähigkeit, relevante Informationen zu filtern und Unwichtiges auszublenden — ein Mechanismus, der auch als selektive Aufmerksamkeit bezeichnet wird (vgl. Abschnitt 2.1 Menschliche Faktoren). Die zentrale Frage im Zusammenhang mit dem Restorff-Effekt und dem Design von Produkten lautet: Wie kann die Aufmerksamkeit der Benutzer gezielt gelenkt werden? Die grosse Herausforderung besteht darin, die Aufmerksamkeit so zu steuern, dass Nutzer sich auf ihre eigentlichen Ziele konzentrieren können und gezielt bei deren Erreichung unterstützt und nicht unnötig abgelenkt werden.

Es gibt verschiedene Ansätze, um Nutzer zum Ziel oder zu den gewünschten Elementen zu führen. Häufig werden visuelle Hervorhebungen eingesetzt. Diese lenken die Aufmerksamkeit gezielt auf bestimmte Elemente — etwa eine Schaltfläche oder eine zentrale Information. Dabei sollte jedoch berücksichtigt werden, dass mehrere visuell hervorgehobene Elemente miteinander konkurrieren: Je mehr Objekte gleichzeitig um die Aufmerksamkeit ringen, desto schwieriger wird es, die Fokussierung auf die tatsächlich relevanten Inhalte oder Informationen zu lenken und umso höher fällt die kognitive Belastung aus. Daraus ergibt sich, dass Gestaltungsmittel wie Farbe, Größe und Kontraste eine wichtige Rolle spielen und deshalb bewusst eingesetzt werden sollten. Durch gezielte Farb- und Kontrasteffekte lassen sich Funktionen visuell unterscheiden und die Lesbarkeit verbessern. Eine klare optische Trennung hilft Nutzerinnen und Nutzern, die verfügbaren Funktionen besser zu erkennen, und kann Fehlbedienungen vorbeugen, etwa indem sie verhindert, dass versehentlich eine falsche Aktion gewählt wird. Eine Veranschaulichung des Restorff-Effekts ist im Anhang C unter dem Abschnitt Restorff-Effekt in Abbildung A.4 dargestellt.

2.3.2 Usability-Prinzipien / UX-Prinzipien

Im folgenden Abschnitt werden weitere grundlegende Prinzipien der Designgestaltung vorgestellt. Dabei handelt es sich weniger um feste Gesetze als vielmehr um praxisorientierte Richtlinien. Diese Richtlinien sind wichtige Werkzeuge, die dazu beitragen, dass Nutzer die Funktionalitäten eines Systems oder Produkts mit möglichst geringem kognitiven Aufwand erfassen und verstehen können. Im Verlauf dieses Abschnitts wird klar, dass sich diese Prinzipien an einigen Stellen überschneiden und sich gegenseitig bestärken oder auch relativieren können.

Erlernbarkeit

Der erste Grundsatz betrifft die sogenannte Erlernbarkeit. Ein System und seine Funktionsweise sollten so gestaltet oder angeordnet sein, dass sie sich leicht erlernen und verstehen lassen. Ein leicht erlernbares System ermöglicht es den Nutzern, die bereitgestellten Funktionen wie etwa Schaltflächen, Menüs usw. korrekt, effizient, effektiv und mit möglichst geringem Erinnerungsaufwand zu nutzen, um Aufgaben zu erledigen oder ihre Ziele zu erreichen, nachdem sie das System einige Male verwendet oder zur Problemlösung eingesetzt haben [13]. Die Erlernbarkeit wird gefördert, wenn das System seine Funktionalitäten schrittweise erkennbar macht, damit sich die Nutzer nicht alles auf einmal merken müssen, wie das Produkt funktioniert. Ein gut erlernbares System weist eine starke Selbstbeschreibung auf, ist also im Grundsatz ohne Hilfe oder Dokumentationen nutzbar. Werden ausserdem die mentalen Modelle der Nutzer oder der Zielgruppe berücksichtigt, trägt dies ebenfalls gezielt dazu bei, die Erlernbarkeit eines Systems zu unterstützen oder zu verbessern, da Nutzer in diesem Fall Wissen aus früheren Erfahrungen direkt anwenden können.

Einprägsamkeit

Der nächste Grundsatz ist die sogenannte Einprägsamkeit. Darunter wird die Fähigkeit der Nutzer verstanden, sich nach einer Phase ohne der System- oder Produktnutzung wieder an die Funktionsweise des Systems oder Produktes zu erinnern. Ein einprägsames System ermöglicht es, nach einer Unterbrechung die verfügbaren Funktionen ohne eine erneute Einarbeitung korrekt, effektiv und effizient zu nutzen. Die Einprägsamkeit ist für Systeme oder Produkte, welche nur selten oder gelegentlich verwendet werden, eine wertvolle und wichtige Eigenschaft.

Werden die Erlernbarkeit und die Einprägsamkeit verglichen, fällt möglicherweise auf, dass sich diese beiden Konzepte teils überschneiden können. Ein System, das sich leicht erlernen lässt, weist häufig auch eine gute Einprägsamkeit auf, da für die Erlernbarkeit klare und leicht erkennbare Muster und Strukturen eingesetzt werden. Umgekehrt kann ein System, das vertraute und bekannte Abläufe einsetzt, sowohl das Erlernen als auch das spätere Erinnern unterstützen. Dennoch unterscheiden sich die beiden Konzepte darin, dass sich die Erlernbarkeit auf die anfängliche Aneignung der Systemnutzung bezieht, während sich die Einprägsamkeit auf die Nachhaltigkeit dieses Wissens nach einer Nutzungspause bezieht. Dabei spielt auch eine Rolle, dass Nutzer nach längeren Nutzungspausen in der Regel nicht dazu bereit sind, lange Anleitungen oder Hilfestellungen zu lesen, um sich wieder mit der Bedienung des Systems vertraut zu machen. Dieses Verhalten steht im Zusammenhang mit dem sogenannten Paradoxon des aktiven Nutzers, das in Kapitel 2 unter

Abschnitt 2.2 beschrieben wurde.

Kontrollierbar und Steuerbarkeit

Als nächstes geht es um den Grundsatz der Kontrollierbarkeit oder Steuerbarkeit oder als das Prinzip der Kontrolle genannt. Dieses Prinzip verlangt, dass die Benutzer eines Produktes oder Systems stets die Kontrolle über die Aktionen und Handlungen des Produktes haben. Eine gute Kontrollierbarkeit erlaubt den Benutzern also, das Produkt oder Handlungsabfolgen während Interaktionsprozessen gewissermassen zu steuern [14].

Ein wichtiger Aspekt dieses Prinzips ist die Möglichkeit, Handlungssequenzen zu unterbrechen und wieder aufzunehmen. Formulare, zum Beispiel Buchungformulare, sind oft oder teilweise in unterschiedliche Eingabekategorien unterteilt. In jedem Schritt ist nur eine Eingabekategorie sichtbar und das Formular wird sodann schrittweise ausgefüllt. Zum Beispiel könnte das Formular in Personeninformationen, Unterkunftsinformationen und Abreise- und Ankunftsdatum unterteilt sein. Ein Nutzer würde dann z.B. seine Personalien, dann Informationen zur Unterkunft und schliesslich in einem letzten Schritt das gewünschte Abreise- und Ankunftsdatum eintippen, wobei er aber selber entscheidet, wann er welche Kategorie abschliesst. Über die Links oberhalb des Formulars kann der Nutzer von einer Kategorie zur anderen Kategorie springen und seine Eingaben allenfalls anpassen (vgl. Anhang A, Abbildung A.5). Auf diese Weise kann der Nutzer Handlungssequenzen unterbrechen, an einer anderen Stelle weitermachen oder eine unterbrochene Sequenz wieder aufnehmen. Dieses Beispiel beschreibt einen weiteren wichtigen Bestandteil, nämlich die Steuerbarkeit der individuellen Interaktionsgeschwindigkeit. Mithilfe der erwähnten Links oberhalb des Formulars, können Nutzer selbst entscheiden, wann und wie schnell sie jeden Interaktionsschritt abschliessen.

Ein weiterer Bestandteil des Prinzips der Kontrolle verlangt, dass Aktionen und Handlungen des Systems umkehrbar sind. Die Möglichkeit, Aktionen rückgängig machen zu können, unterstützt die Nutzer, Fehler zu beheben und das System ungehemmt zu nutzen und zu erforschen [15]. Gleichzeitig werden die Benutzer ermutigt, Aktionen oder Handlungsabläufe, welche sie (noch) nicht kennen, auszuprobieren.

Flexibilität

Ein weiterer Gestaltungsgrundsatz ist die Flexibilität. Ein flexibles System ermöglicht es den Endnutzern, Funktionen entweder bereits zu Beginn oder fortlaufend während des Interaktionsprozesses anzupassen oder hinzuzufügen [16]. Wenn ein System an die Bedürfnisse der Nutzer angepasst werden kann, so stärkt dies einerseits das

Gefühl von Kontrolle und andererseits wird die Zielerreichung unterstützt, da die Nutzer selbst entscheiden können, welche Funktionen sie benötigen und welche nicht. Das heisst, die Nutzer haben freie Hand beim Entscheiden, wie viel Komplexität ein System oder ein Prozess innerhalb eines Produktes aufweist. Eine Illustration findet sich in Anhang A in Abbildung A.6. Das Mass an Flexibilität hängt stark von der Zielgruppe und deren Bedürfnissen, Zielen und Fähigkeiten ab. Nutzer mit viel Erfahrung mit dem Produkt haben in der Regel andere Ansprüche an die Flexibilität als Nutzer mit wenig Erfahrung. Erstere Gruppe wird in der HCI oft als Experten bezeichnet. Diese legen häufig Wert darauf, ihre Ziele möglichst effizient und effektiv zu erreichen. Eine hilfreiche Form der Flexibilität in diesem Fall wäre beispielsweise die Möglichkeit, wiederkehrende Abläufe zu automatisieren oder Aktionen und Prozesse direkt über eine Kommandozeile ausführen zu können.

Für weniger erfahrene oder neue Nutzer sind hingegen andere Aspekte der Flexibilität von Bedeutung. Diese Gruppe profitiert von modifizierbaren Schaltflächen, die sich in ihrer Komplexität reduzieren oder leicht erweitern lassen [17]. Beispielsweise können Systeme, die zwischen einem Einsteigermodus und Expertenmodus unterscheiden, dazu beitragen, mentale Überforderung zu vermeiden, indem in einem ersten Schritt nur die relevanten Funktionen sichtbar sind. Auch assistierende Funktionen, wie geführte Abläufe oder im Kontext eingebettete Hilfestellungen, stellen eine Form der Flexibilität dar, welche unerfahrenen Nutzern hilft, ihre Ziele zu erreichen [18]. Damit zeigt sich, dass sich Flexibilität in diversen Formen umsetzen lässt, also nicht nur aus Anpassungsmöglichkeiten besteht, sondern auch in der Fähigkeit eines Systems, sich an unterschiedliche Erfahrungsniveaus der Nutzer anzupassen.

Fehlertolerant

Ein fehlertolerantes System reagiert flexibel auf Fehleingaben und liefert bei Bedarf verständliche Rückmeldungen (Feedback), die Nutzerinnen und Nutzer dabei unterstützen, Fehler zu beheben oder künftig zu vermeiden. Ein klassisches Beispiel hierfür ist das Ausfüllen eines Formulars in einer Webschnittstelle. Um die kognitive Belastung möglichst gering zu halten, ist es sinnvoll, wenn das System Fehleingaben automatisch erkennt und eine kurze und klar formulierte Fehlermeldung in unmittelbarer Nähe des entsprechenden Eingabefelds erzeugt. Diese Rückmeldung sollte idealerweise nicht erst beim Absenden des Formulars erfolgen. Andernfalls müsste die Person zunächst die Fehlermeldung lokalisieren und anschliessend die Eingabe korrigieren, was ein zusätzlicher kognitiver Aufwand darstellt. Wird die Fehlermeldung hingegen bereits angezeigt, während sich der Mauszeiger noch im betreffenden Eingabefeld befindet, ist davon auszugehen, dass auch die Aufmerksamkeit des Nut-

zers noch auf dieses Feld gerichtet ist. Die Rückmeldung wird somit schneller wahrgenommen und kann rascher verarbeitet werden. Ein solches Verhalten des Systems wird in der HCI als modeless Feedback bezeichnet.

Konsistenz

Beim Gestaltungsgrundsatz der Konsistenz existieren grundsätzlich zwei Aspekte. Im Folgenden wird nur ein Aspekt, der der externen Konsistenz, welcher für die Anpassung von Produktdesigns an die mentalen Modelle der Zielbenutzer wichtig ist, beschrieben.

Die externe Konsistenz geht über die Grenzen eines einzelnen Systems hinaus. Nach [19] beschreibt sie die Übereinstimmung zwischen einer Anwendung und anderen, vergleichbaren Produkten. Wenn eine Anwendung in ihrer Darstellung, Kommunikation und ihrem Verhalten auf bekannte Konventionen zurückgreift, die auch in anderen Anwendungen verwendet werden, verhält sie sich extern konsistent. Eine Möglichkeit, externe Konsistenz zu erreichen, besteht darin, Abläufe, Schaltflächen, Bezeichnungen und andere Funktionalitäten so zu gestalten, dass sie für Nutzer mit Erfahrungen aus anderen Systemen wiedererkennbar und vergleichbar sind [1]. An dieser Stelle lässt sich festhalten, dass externe Konsistenz damit die Anpassung des Designs an die mentalen Modelle der Nutzer erleichtert, was wiederum die Erlernbarkeit und Einprägsamkeit eines Systems fördert.

Zur Veranschaulichung kann folgendes Beispiel betrachtet werden: Nutzt eine Software auf einem Desktop-Computer für Speicherfunktionen immer dasselbe Symbol und dieselbe Schaltfläche, weist sie interne Konsistenz auf. Verwendet diese Software darüber hinaus dieselben Bedienelemente wie das Betriebssystem selbst, ist sie auch extern konsistent.

2.3.3 Weitere Gestaltungselemente

Angebotscharakter

Der Angebotscharakter (Affordances) beschreibt die möglichen Handlungen oder Aktionen, die ein Objekt oder Element durch seine Gestaltung oder Erscheinung signalisiert [20]. Manchmal spricht man auch von der Beziehung zwischen Elementen und Benutzern im Hinblick auf die möglichen Aktionen. Die Form, Darstellung und Visualisierung eines Objekts beeinflussen massgeblich, wie diese Handlungsmöglichkeiten von Nutzern wahrgenommen und interpretiert werden. Die Hinweise zur Art der Nutzung eines Objektes oder Elementes sind je nach Erscheinungsbild des Objektes unterschiedlich stark. Ein klassisches Beispiel ist eine Schaltfläche in einer Webanwendung, etwa ein runder, am Rand schattierter Button, der visuell

als klickbares Element erkennbar ist (vgl. Anhang A, Abbildung A.7). Ein starker Angebotscharakter kann helfen, ein Design selbstbeschreibender zu gestalten, also dem Nutzer mitzuteilen, welche Funktionalität das System bereithält.

Feedback

Ein Feedback sollte immer dann erzeugt werden, wenn ein System oder Produkt dem Nutzer eine Zustandsänderung oder andere Informationen – etwa das Ergebnis einer Aktion – übermittelt. Dabei ist es wichtig, dass das Feedback unmittelbar nach der jeweiligen Aktion erfolgt. Verzögertes Feedback kann bei Nutzern zu Unsicherheit oder Verwirrung führen. Zudem sollte ein Feedback stets informativ sein, sodass erkennbar ist, was geschehen ist und – falls notwendig – wie weiter vorzugehen ist [21].

Die Herausforderung bei der Gestaltung von Feedback liegt darin, den richtigen Zeitpunkt, die angemessene Menge und die geeignete Form zu wählen. Zu häufiges oder aufdringliches Feedback, insbesondere solches, das aktiv bestätigt werden muss, kann zu Ablenkungen führen und die kognitive Belastung erhöhen. Ein typisches Beispiel hierfür sind Dialogfenster, die geschlossen werden müssen, bevor die Interaktion fortgesetzt werden kann. Werden solche Rückmeldungen zu häufig eingesetzt, besteht die Gefahr, dass sie von den Nutzern künftig ignoriert oder reflexartig geschlossen werden, selbst dann, wenn sie relevante Informationen enthalten. Dadurch steigt das Risiko, dass wichtige Hinweise übersehen oder verpasst werden.

Aus diesem Grund sollte die Gestaltung von Feedback strukturiert und gut geplant erfolgen. Grundsätzlich ist es hilfreich, sich an folgenden Prinzipien zu orientieren: Ist die Information nicht dringend oder kritisch, sollte das Feedback unaufdringlich erfolgen – idealerweise in einer Form, die keine Handlung durch die Nutzer verlangt. Müssen hingegen wichtige Informationen vermittelt werden, sollte das Feedback so gestaltet sein, dass es die Aufmerksamkeit gezielt auf sich zieht. Bestätigungsdialoge können in solchen Fällen sinnvoll sein, sollten aber mit Bedacht eingesetzt werden. Selbst bei relevanten Informationen ist nicht immer eine sofortige Reaktion notwendig. Ausnahmen bestehen bei potenziell schwerwiegenden Aktionen, die nur mit grosser Mühe rückgängig gemacht werden können. In solchen Fällen kann eine bewusste Nutzerbestätigung sinnvoll sein. Zusammenfassend lässt sich sagen, dass Feedback entsprechend der Relevanz und Dringlichkeit der übermittelten Informationen priorisiert werden sollte [20].

Kapitel 3

Messen der Usability und User Experience

Dieses Kapitel beginnt mit einem Vergleich verschiedener Ansätze zur Durchführung von Usability- und User Experience-Studien, welche im Kontext der HCI häufig auch als Usability- oder UX-Tests bezeichnet werden. Anschliessend werden verschiedene Metriken vorgestellt, die zur Erhebung und Bewertung von Usability und der User Experience verwendet werden können. Die Betrachtung erfolgt zunächst vor allem konzeptuell und unabhängig von den im Rahmen dieser Arbeit eingesetzten Evaluationsmethoden. Es handelt sich dabei um grundlegende Kenngrössen oder Messwerte, deren konkrete Manifestation durch Variablen abhängig vom Anwendungskontext erfolgt.

3.1 Usability-Tests

Nachfolgend werden zwei weitverbreitete und häufig eingesetzte Durchführungsmethoden von Usability- und User Experience-Studien, auch als Usability-Tests bezeichnet, erläutert. Die in diesem Abschnitt zusammengefassten Usability- und User Experience-Studien und weitere Studienansätze werden ausführlich in [22] und [23] beschrieben.

Moderierte Usability-Tests

Eine weit verbreitete und häufig eingesetzte Methode zur Bewertung der Usability und User Experience von Produkten oder Systemen ist der klassisch moderierte Usability-Test. Diese Tests werden in der Regel mit etwa 12 bis 15 Teilnehmenden durchgeführt, wobei die Anzahl an Teilnehmenden stark variieren kann. Grundsätzlich gilt: Je grösser die Stichprobe, desto aussagekräftiger sind die gewonnenen Daten über die Usability und User Experience des Produkts.

Diese Art von Test findet typischerweise im Rahmen einer Eins-zu-eins-Sitzung zwischen einer Testperson und einer Testleitung bzw. einem UX-Spezialisten in einem geschützten Umfeld (Usability-Labor) statt. Den Teilnehmenden wird eine festgelegte Aufgabenliste vorgegeben, die sie selbstständig bearbeiten sollen. Während der Aufgabenbearbeitung wird häufig die Think-Aloud-Methode eingesetzt [24]. Dabei spricht die Testperson ihre Gedanken während der Bearbeitung laut aus, damit der Testleiter das Vorgehen nachvollziehen kann. Bei Rückfragen oder auftretenden Schwierigkeiten steht die Testleitung unterstützend zur Verfügung.

Das Vorgehen der Testperson wird systematisch beobachtet und dokumentiert. Neben qualitativen Beobachtungen werden auch spezifische Daten wie Performance-Messwerte, Bearbeitungszeiten oder Interaktionsabbrüche erfasst. Zusätzlich werden während des Tests Usability-Probleme (Issues), Verständnisfragen oder Unklarheiten notiert, um daraus Rückschlüsse für die Weiterentwicklung oder Verbesserung des Produkts ziehen zu können.

Diese Art von Usability-Test ist besonders hilfreich, um Usability-Probleme (Usability-Issues) zu entdecken und deren Ausmass zu beurteilen. Zudem können durch die Think-Aloud-Methode tiefere Erkenntnisse gewonnen werden, wie Nutzer ein Design wahrnehmen und nutzen und wie ein Design auf Nutzer wirkt.

Unmoderierte Usability-Tests

Im Gegensatz zu den zuvor beschriebenen moderierten Usability-Tests kommen heute zunehmend auch unmoderierte Usability-Tests zum Einsatz, um die Usability eines Produkts und die damit verbundene User Experience zu evaluieren. Bei dieser Testform werden die Teilnehmenden nicht in Echtzeit von einer Testleitung oder UX-Fachperson begleitet oder beobachtet. Das bietet den Vorteil, dass diese Art von Tests mit weitaus mehr Teilnehmern durchgeführt werden kann, als moderierte Usability-Tests.

Die Probanden erhalten auch hier eine Liste mit vorab definierten Aufgaben, die anschliessend selbstständig gelöst werden sollen. Wo die Probanden den Test durchführen, können diese in der Regel selbst bestimmen. Alle zur Evaluierung der Usability oder User Experience benötigten Daten werden automatisch gesammelt. Dies bietet den Vorteil, dass weitaus mehr Daten als bei den moderierten Tests gesammelt werden können. Die Schwierigkeit hier ist jedoch, dass im Vorfeld geplant werden muss, welche Daten zur Beurteilung der Usability und User Experience gesammelt werden sollen. Dabei kann es herausfordernd sein, vorab die korrekten Daten zu bestimmen, welche dann tatsächlich eine Beurteilung der Usability und der gesamten User Experience zulassen.

Da die Teilnehmer aber nicht direkt beobachtet werden können, müssen Proble-

me, Usability-Issues oder auffälliges Verhalten im Nachgang durch die gesammelten Daten rekonstruiert werden, was teils keine leichte Aufgabe ist, da Daten lückenhaft, unvollständig oder schlicht nicht aussagekräftig genug sein könnten. Es ist bei dieser Art von Tests also zeitaufwändiger, tiefere Einblicke in die Präferenzen, Motivationen oder Verhaltensmuster der Nutzer zu erhalten.

3.2 Usability und User Experience-Metriken

Damit die Usability mitsamt der User Experience einer Nutzerinteraktion beurteilt werden können, werden Methoden benötigt, um diese beiden Aspekte zu messen. Hierfür können unterschiedliche Usability- und User Experience-Metriken (UX-Metriken) eingesetzt werden.

In welcher Form sich diese Metriken ausprägen und manifestieren, hängt stark vom Kontext, also vom System oder Produkt, oder aber auch von der Untersuchung, welche im Rahmen der Interaktion durchgeführt werden soll, ab. Die Definition der Metriken hängt sodann davon ab, in welcher Form das Nutzerverhalten während der Interaktion untersucht werden soll. Falls bereits Daten existieren, so können auch auf diese bereits bestehenden Daten zurückgegriffen werden, um Metriken auszuwählen [25].

Eine mögliche Kategorisierung der Metriken zur Beurteilung der durch ein Produkt erzeugten Usability und User Experience könnte wie folgt aussehen:

- Performance-Metriken
- Issues-Based-Metriken (Usability-Problem orientierte Metriken)
- Self-reported-Metriken
- Kombinierte und vergleichende Metriken

Jede der oben genannten Kategorien soll nachfolgend kurz beschrieben werden. Jede beschriebene Metrik wird in [22] und [23] sehr detailliert anhand Beispielen vorgestellt.

3.2.1 Performance-Metriken

Der im Rahmen dieser Arbeit durchgeführte Usability-Test verfolgt unter anderem das Ziel, sowohl das Nutzerverhalten als auch die Nutzer-Performance zu erfassen. Zur Messung der Effektivität und Effizienz eines digitalen Produkts sowie zur Analyse des Nutzerverhaltens kommen sogenannte Performance-Metriken zum Einsatz.

Die Performance-Metriken stellen abhängige Variablen dar, anhand derer menschliches Verhalten gemessen wird [5]. Diese Metriken ermöglichen es, das Verhalten von Nutzern zu messen und auf Basis vordefinierter Kriterien zu bewerten. Welche Metriken konkret verwendet werden, hängt davon ab, welche Aspekte des Nutzerverhaltens im Fokus der Untersuchung stehen. Eine zentrale Voraussetzung für die Anwendung von Performance-Metriken ist die Zuweisung klar definierter Aufgaben: Der Start- und Zielzustand der jeweiligen Aufgabe müssen eindeutig festgelegt und den Teilnehmenden im Vorfeld kommuniziert werden. Zu den Performance-Metriken, die auch in der vorliegenden Studie Anwendung finden, zählen unter anderem:

- Aufgaben-Erfolg
- Auftragszeit
- Erlernbarkeit
- Einprägsamkeit
- Konsistenz
- Fehler
- Effizienz

Aufgaben-Erfolg

Die Aufgaben-Erfolg-Metrik ist eine einfache Metrik, um zu bestimmen, ob eine Aufgabe oder ein Problem erfolgreich oder nicht erfolgreich gelöst wurde. Um dies bestimmen zu können, braucht eine Aufgabe ein vorab definiertes Ziel oder einen Endzustand sowie einen Anfangszustand. Nach Abschluss einer Aufgabe wird dann anhand des vorab definierten Zielzustandes bestimmt, ob die Aufgabe erfolgreich abgeschlossen wurde oder nicht. Eine Aufgabe muss dabei nicht zwangsläufig aus mehreren Schritten oder Teilzielen bestehen, sondern kann auch beliebig fein granular definiert sein. Ein einziger Mausklick auf eine Schaltfläche könnte zum Beispiel ebenfalls eine Aufgabe sein.

Auftragszeit

Als Auftragszeit wird die Zeitspanne bezeichnet, die ein Nutzer benötigt, um eine Aufgabe bis zu einem vorab definierten Endzustand zu bearbeiten. Dabei ist es irrelevant, ob die Aufgabe erfolgreich oder nicht erfolgreich gelöst wurde – entscheidend ist allein, dass der definierte Zielzustand erreicht wurde, unabhängig vom

Ergebnis. Möglich ist es aber, eine maximale Auftragszeit zu definieren, bis zu der ein Endzustand erreicht werden muss. Wurde der Endzustand bis zu einer maximal festgelegten Zeit nicht erreicht, so kann die Aufgabe als nicht gelöst betrachtet werden. Ein Beispiel für die Auftragszeit ist die Zeit, die benötigt wird, um ein Produkt in den Warenkorb zu legen. Der Endzustand ist in diesem Fall erreicht, wenn sich mindestens ein Produkt im Warenkorb befindet.

Fehler

Es gibt viele verschiedene Ausprägungen von Fehlern. Die Art und Weise, wie Fehler definiert oder bestimmt werden, hängt stark vom Kontext oder von der Aufgabe ab. Je nach eingesetztem Produkt und Aufgabe können Fehler in unterschiedlichen Formen auftreten. Aus einem sehr allgemeinen Blickwinkel betrachtet, kann ein Fehler wie folgt beschrieben werden: Ein Fehler ist eine Handlung oder Aktion, welche einen Nutzer daran hindert, sein Ziel mit höchster Effizienz und Effektivität zu erreichen. Die wichtigsten Fehlertypen, die für diese Arbeit relevant sind und während der Durchführung der User Experience wahrscheinlich am häufigsten auftreten können, sind Fehlclicks. In dieser Arbeit werden zwei Arten von Fehlclicks unterschieden:

Die erste Art umfasst Clicks, die ein falsches Zielelement betreffen und unmittelbar zu einer fehlerhaften Systemaktion oder einem unerwünschten Zustand führen. Ein typisches Beispiel wäre das versehentliche Klicken auf die Lösch-Schaltfläche, statt auf die Archivieren-Schaltfläche in einer E-Mail-Anwendung, wodurch eine E-Mail irrtümlich gelöscht wird.

Die zweite Art beschreibt Clicks, die zwar nicht direkt zu einem falschen Systemzustand führen, den Nutzer jedoch in seiner Zielverfolgung nicht voranbringen. Ein Beispiel hierfür ist die Navigation in einem Menü mit vertikal angeordneten Produktkategorien. Klickt der Nutzer versehentlich auf eine nicht gewünschte Kategorie, werden die entsprechenden Unterkategorien angezeigt, allerdings nicht die, die tatsächlich relevant oder gewollt sind. Dieser Klick führt das System nicht in einen fehlerhaften Zustand, bringt den Nutzer aber auch nicht näher an sein Ziel. Der Interaktionsprozess verlängert sich dadurch unnötig, was zu einer nicht optimalen Effizienz führt.

Erlernbarkeit

Der Begriff der Erlernbarkeit wurde bereits in Kapitel 2, Abschnitt 2.3.2, als grundlegendes UX-Prinzip erläutert. Darüber hinaus wird Erlernbarkeit häufig auch als Metrik verwendet, um die Leistungsfähigkeit eines Lernprozesses während der In-

teraktion mit einem System zu erfassen.

Aus einer allgemeinen Perspektive beschreibt Erlernbarkeit das Ausmass, in dem ein Nutzer neue Fähigkeiten erlernen kann, um ein Produkt effektiv und effizient zu nutzen. Dieses Ausmass lässt sich unter anderem anhand der benötigten Zeit und des erforderlichen Aufwands messen, welche Nutzer investieren müssen, um das Produkt in einer sehr effektiven und effizienten Weise zu nutzen. Die Metrik Erlernbarkeit spielt eine zentrale Rolle bei der Bewertung der Usability und der User Experience eines Produkts oder eines Designs. Eine steile Lernkurve deutet in der Regel darauf hin, dass Nutzer dazu in der Lage sind, das System im Verlauf der Zeit effizient und präzise zu nutzen und ihre Ziele zuverlässiger zu erreichen. Der Lernprozess selbst kann sich über kürzere oder längere Zeiträume erstrecken. Die Geschwindigkeit, mit der Nutzer lernen, hängt dabei sowohl von ihren Vorerfahrungen als auch von den Erfahrungen ab, die sie im Verlauf der Interaktion sammeln.

Zur Messung der Erlernbarkeit gibt es unterschiedliche Ansätze. In der Praxis werden häufig die benötigte Zeit und der Aufwand herangezogen, bis ein bestimmtes Mass an Effizienz und Effektivität bei der Nutzung erreicht ist. Die Erlernbarkeit liefert somit wertvolle Hinweise darauf, wie schnell und effizient Nutzer Kompetenzen im Umgang mit einem System entwickeln können.

Einprägsamkeit

Einprägsamkeit kann ebenfalls als Performance-Metrik eingesetzt werden. Sie gibt Aufschluss darüber, in welchem Mass Nutzer die Funktionsweise eines Systems oder Produktes einprägen können. In vielen Interaktionen greifen Personen wiederholt auf bestimmte Funktionen oder Abläufe innerhalb eines Systems zurück. Das bedeutet, dass sich bestimmte Interaktionsprozesse, etwa die Navigation innerhalb einer Webanwendung, im Verlauf der Nutzung mehrfach wiederholen.

Das Mass an Einprägsamkeit lässt sich dadurch erfassen, dass die Zeit oder der Aufwand gemessen wird, welche Nutzerinnen und Nutzer für diese mehrfach ausgeführten Handlungen im Verlauf der Interaktion aufbringen müssen. Eine hohe Einprägsamkeit zeigt sich darin, dass solche wiederholten Abläufe mit zunehmender Nutzung schneller und präziser und mit weniger Fehlern ausgeführt werden können. Die Einprägsamkeit ist daher ebenfalls eine hilfreiche Metrik, um eine User Experience zu messen und anschliessend zu beurteilen.

Effizienz

Zur Beurteilung der Effizienz wird der erforderliche Aufwand betrachtet, den Nutzerinnen und Nutzer aufbringen müssen, um eine Aufgabe erfolgreich abzuschliessen.

Dieser Aufwand lässt sich in der Regel anhand der Anzahl notwendiger Aktionen oder Interaktionen messen. Eine solche Aktion kann unterschiedliche Formen annehmen, beispielsweise das Auswählen eines Links oder das Klicken auf eine Schaltfläche. Jede dieser Aktionen erfordert einen gewissen Aufwand seitens der Nutzenden. Grundsätzlich gilt: Je mehr Aktionen zur Erfüllung einer Aufgabe notwendig sind, desto höher ist der Gesamtaufwand und desto geringer ist typischerweise die Effizienz. Dabei wird der Aufwand häufig in zwei Kategorien unterteilt: kognitiver Aufwand und physischer Aufwand.

- Kognitiver Aufwand umfasst mentale Prozesse wie das Erkennen relevanter Elemente auf der Benutzeroberfläche, das Treffen von Entscheidungen über die nächsten Schritte oder die Interpretation von Systemrückmeldungen.
- Physischer Aufwand bezieht sich auf die tatsächlichen Bewegungen, die zur Ausführung einer Aktion erforderlich sind, beispielsweise Mausbewegungen oder Tastatureingaben.

In einigen Fällen wird Effizienz auch als Verhältnis zwischen der Aufgabenerfüllungsrate und der dafür benötigten Bearbeitungszeit definiert.

3.2.2 Self-reported-Metriken

Im Gegensatz zu Performance-Metriken geben sogenannte Self-Reported-Metriken mehr Aufschluss darüber, wie die Nutzer ein System oder das Design eines Systems persönlich wahrnehmen und damit interagieren. Self-Reported-Metriken helfen, Informationen darüber zu sammeln, wie Nutzer über ein Produkt denken oder fühlen. Die Art und Weise, wie ein Nutzer über ein Produkt denkt oder fühlt, kann ebenso wichtig sein wie das Wissen, dass Nutzer mehr Zeit oder mehr Mausklicks benötigen, um beispielsweise ein Objekt zu finden. Ein ausschlaggebender Grund, weshalb ein Nutzer ein Produkt nutzt, liegt oftmals im Umstand, dass das Produkt eine User Experience bietet, welche den Nutzer glücklich macht oder positiv stimmt. Die Messungen und Beurteilungen, welche mit dieser Art von Metriken durchgeführt werden, werden in der Regel durch die Nutzer selbst vorgenommen, zum Beispiel durch Umfragen. Einige dieser Metriken werden nun nachfolgend kurz zusammengefasst.

Bewertungsskalen

In Bewertungsskalen werden häufig Aussagen zur Zustimmung oder Ablehnung zu Inhalten genutzt, um Einschätzungen von Nutzern zu erfassen. Eine der gebräuchlichsten Formen ist die Likert-Skala [22], bei der Teilnehmende angeben, in

welchem Masse sie einer Aussage zustimmen – typischerweise auf einer Skala von „stimme überhaupt nicht zu“ bis „stimme voll und ganz zu“.

Diese Zustimmungsskalen können sowohl verbal als auch numerisch sein. Eine numerische Variante verwendet beispielsweise Ziffern von 1 bis 5 oder von 1 bis 7, wobei 1 für „stimme überhaupt nicht zu“ und 5 bzw. 7 für „stimme voll und ganz zu“ steht. Je nach Fragestellung und Kontext werden unterschiedliche Wertebereiche eingesetzt – etwa mit fünf, sieben oder teils zehn Ziffern, wobei die Wahl der Skalenlänge unter anderem von der gewünschten Differenzierung der Befragten abhängt. Solche Skalen ermöglichen es, qualitative Daten zu sammeln und eignen sich besonders gut zur Messung von Aspekten wie Zufriedenheit, Vertrauen, wahrgenommener Benutzerfreundlichkeit oder dem Gefühl der Kontrolle während der Interaktion.

Nachbefragung

Nachbefragungen liefern wertvolle Daten und Erkenntnisse, insbesondere wenn das Interesse besteht, die subjektive Einschätzung der Nutzenden zur Usability und User Experience nach der Interaktion zu erfassen. Die Teilnehmenden erhalten hierbei die Möglichkeit, offen ihre Einschätzung zu äussern – beispielsweise welche Funktionalitäten sie als besonders positiv oder negativ empfunden haben. Ausserdem können sie mitteilen, welche Aspekte der Produkte intuitiv verständlich waren und bei welchen sie mehr Zeit oder Aufwand investieren mussten, um diese zu verstehen. Auf Grundlage dieser Rückmeldungen lassen sich Bereiche identifizieren, die von den Nutzenden als schwer verständlich wahrgenommen wurden, also solche, welche zu einem Usability-Issue führen könnten. Diese Bereiche können anschliessend im Rahmen eines iterativen Entwicklungsprozesses überarbeitet und stärker an die Bedürfnisse und Erwartungen (mentale Modelle) der Zielgruppe angepasst werden.

Für solche Nachbefragungen gibt es einen hierfür von John Lewis 1991 [26] entwickelten Fragebogen, den sogenannten After-Score-Questionnaire (ASQ). Dieser wird vor allem eingesetzt, nachdem Teilnehmer eine Menge an Aufgaben abgeschlossen haben und die Einschätzung der einzelnen Aufgabenschwierigkeiten untersucht werden soll.

Um die Nutzer nach der Wahrnehmung der gesamten User Experience zu befragen, gibt es mehrere entwickelte Fragebögen. Ein weit verbreiteter ist der sogenannte System Usability Scale (SUS), der ursprünglich von John Brooke 1996 entwickelt wurde [27]. Dieser Fragebogen enthält zehn Aussagen, zu denen die Nutzer ihr Ausmass an Zustimmung oder Ablehnung bekanntgeben. Zu jeder Aussage können die Nutzer auf einer Skala von 1 - 5 ihre Zustimmung oder Ablehnung bewerten.

Mithilfe der Antworten wird dann ein Gesamtscore zwischen 0 und 100 berechnet, wobei 0 für eine sehr schlechte User Experience steht und 100 für eine optimale User Experience.

3.2.3 Problem-basierte Metriken (Usability-Issues)

Ein Usability-Issue wird meist anhand eines auffälligen oder vom normalen Verhalten abweichenden Interaktionsverhaltens, wie Verwirrung, Frustration, Übersehen von Inhalten, Fehlinterpretationen oder schlechter Performance, identifiziert. Usability-Issues müssen nicht immer negativ sein. Es gibt auch positive Usability-Issues, in diesem Fall wird oft auch von Usability-Erkenntnissen (Usability-findings) gesprochen. Ein Usability-finding kann zum Beispiel eine bestimmte Vorgehensweise sein, um eine Aufgabe zu bewältigen.

Bei der Identifikation eines Usability-Issue sind also mindestens die folgenden zwei Aspekte vertreten: Interpretation des Nutzerverhaltens und das tatsächliche Verhalten (effektive Handlungen des Nutzers). Deshalb bieten sich für das Finden von Usability-Issues die in diesem Kapitel unter Abschnitt 3.1 beschriebenen moderierten Tests an, da diese Art von Tests in der Regel durch eine Eins-zu-Eins-Session erfolgen und daher das Nutzerverhalten direkt beobachtet werden kann. Die Identifikation von Usability-Issues kann helfen, iterativ das Produkt- oder Systemdesign zu verbessern, denn Usability-Issues könnten darauf hinweisen, dass das aktuelle Design besser an die mentalen Modelle der Nutzer angepasst werden könnte.

Tatsächliche Usability-Issues vs. scheinbare Usability-Issues

Bei der Identifikation von Usability-Problemen besteht die Schwierigkeit darin, zwischen tatsächlichen Problemen und solchen zu unterscheiden, die lediglich vom typischen Nutzungsverhalten abweichen und nicht zwangsläufig auf Designmängel hindeuten. Ein möglicher Hinweis auf ein tatsächliches Usability-Problem ist die Häufigkeit, mit der es bei den Testpersonen auftritt.

Wenn beispielsweise zehn von 20 Teilnehmenden in einem Menü wiederholt eine falsche Option auswählen und anschliessend ein gesuchtes Element an einer falschen Stelle innerhalb der Anwendung suchen, deutet dies auf ein möglicherweise systematisches Problem im Design hin. In solchen Fällen ist es sinnvoll, die zugrunde liegenden Gestaltungsentscheidungen zu hinterfragen und gegebenenfalls anzupassen. Tritt ein solches Verhalten jedoch nur bei ein oder zwei Personen auf, sind weiterführende Überlegungen notwendig. Es stellt sich die Frage, ob dieses Verhalten nachvollziehbar und logisch begründet ist.

Lässt sich die Interaktion aus Sicht der betroffenen Personen nachvollziehbar

erklären, etwa aufgrund von Vorerfahrungen, bestimmten Erwartungen oder durch Fehlannahmen, kann dies dennoch auf ein relevantes Usability-Problem hinweisen. Ist das Verhalten hingegen nicht erklärbar, und können die Nutzenden selbst keine Begründung für ihr Vorgehen liefern, liegt die Vermutung nahe, dass es sich um ein eigenartiges Verhalten handelt – also um eine individuelle Abweichung, die nicht auf mangelhafte Usability zurückzuführen ist.

Usability-Issue-Metriken

Sobald Usability-Probleme identifiziert wurden, können diese gemessen werden. Was gemessen werden soll, hängt davon ab, welche Fragen beantwortet werden sollen. Die gängigsten Arten, Usability-Probleme zu messen, sind die folgenden:

- Anzahl eindeutiger Usability-Issues
- Häufigkeit von Usability-Issues pro Teilnehmer
- Anteil der betroffenen Teilnehmer in Prozent
- Issues nach Kategorie

Bei der Anzahl eindeutiger Usability-Issues wird erfasst, wie viele eindeutig identifizierte Usability-Probleme im Verlauf eines Tests auftreten. Dabei wird pro Testdurchlauf gezählt, ob ein bestimmtes Issue beobachtet wurde – mehrfaches Auftreten desselben Issues bei unterschiedlichen Personen wird dabei nicht mehrfach berücksichtigt, sondern nur einmal gezählt.

Der zweite Punkt der oben genannten Liste bezieht sich auf die Gesamtzahl aller Usability-Issues für einen Teilnehmer während eines Testdurchlaufs, wobei auch hier wiederholt auftretende Probleme mehrfach gezählt werden. So wird sichtbar, welche Probleme häufig vorkommen – unabhängig davon, ob sie immer bei denselben oder bei unterschiedlichen Personen auftreten.

Der dritte Punkt beschreibt den Prozentsatz der Testpersonen, bei denen ein bestimmtes, eindeutig identifiziertes Usability-Problem aufgetreten ist.

Wenn Usability-Probleme nach Kategorien gruppiert werden, dient dies dem Ziel, Bereiche des Produktes oder Systems zu identifizieren, die besonders kritisch sind oder als besonders kritisch gewertet werden. Solche Kategorien können zum Beispiel Navigation, Suche, Löschvorgänge oder Ähnliches umfassen. Nach der Einteilung kann es sinnvoll sein, die Häufigkeit der Issues pro Kategorie zu analysieren, um Teile des Produktes oder der Anwendung zu optimieren oder besser an die mentalen Modelle der Nutzer anzupassen.

3.2.4 Kombinierte und vergleichende Metriken

Schliesslich gibt es auch kombinierte und vergleichende Metriken. Dabei handelt es sich in der Regel um Kombinationen verschiedener Performance-Metriken – etwa Erfolgsquote, Bearbeitungszeit oder Anzahl der Fehler – mit Einschätzungen bezüglich der Nutzerbewertungen. Solche Metriken bieten eine umfassendere Perspektive und können helfen, ein ganzheitliches Bild der User Experience zu erhalten.

Eine einfache, aber dennoch aussagekräftige Methode zur Kombination mehrerer Metriken besteht darin, jeweils zwei oder mehr Kennwerte zu einem Datenpunkt zusammenzufassen – zum Beispiel die Kombination aus Bearbeitungszeit und Anzahl der Fehlklicks. Auf Grundlage dieser Datenpunkte lassen sich anschliessend Zielvorgaben definieren. Angenommen, das Ziel besteht darin, dass die Aufgaben in weniger als einer Minute und mit höchstens fünf Fehlklicks abgeschlossen werden. Im Anschluss wird dann untersucht, wie häufig dieses Ziel tatsächlich erreicht wurde. Auf diese Weise lassen sich verschiedene Leistungsindikatoren sinnvoll zusammenführen und bewerten.

Falls keine optimale Zielzeit festgelegt werden kann (was häufig der Fall ist), könnte es sinnvoll sein, die Durchführungszeiten der schnellsten bzw. der langsamsten Teilnehmer zu betrachten und die restlichen gemessenen Durchführungszeiten in Relation zu diesen Zeiten zu betrachten. Um auszudrücken, wie nahe eine gemessene Zeit t_k an der besten (kürzesten) Zeit liegt, kann ein Prozentwert p berechnet werden. Dieser Wert gibt an, zu welchem Anteil die optimale Zeit erreicht wurde – also wie viel Prozent der Weg von der schlechtesten zur besten Zeit geschafft ist. Die Berechnung erfolgt mit folgender Formel:

$$p = \frac{t_{\max} - t_k}{t_{\max} - t_{\min}} \quad (3.1)$$

Dabei gilt:

- t_{\max} : Die längste (schlechteste) Zeit
- t_{\min} : Die kürzeste (beste) Zeit
- N : Die Gesamtanzahl an Zeitmessungen
- t_k : Die beobachtete Zeit bei der k -ten Messung, wobei $1 \leq k \leq N$

Der Wert p liegt zwischen 0 und 1 (0% und 100%). Ein Wert von 1 (bzw. 100%) bedeutet, dass die beste Zeit vollständig erreicht wurde. Ein Wert von 0 (bzw. 0%) bedeutet, dass nur die schlechteste Zeit erzielt wurde. Je höher der Wert, desto näher liegt die beobachtete Zeit an der optimalen Zeit.

Kapitel 4

Durchführung und Evaluation des Usability- und User Experience-Tests

Das vierte und letzte Kapitel dieser Arbeit behandelt mehrere Themen. Es beginnt mit der Vorstellung der gewählten Methode für den Usability- und User Experience-Test, geht dann auf die Umsetzung und den Aufbau der dafür entwickelten Webseite ein und beschreibt schliesslich die durchgeführten Versuchsexperimente sowie deren Ergebnisse im Detail.

Zunächst wird die gewählte Testmethode motiviert und beschrieben. Anschliessend folgt eine Erläuterung des Testaufbaus, also wie der Test aufgebaut ist und abläuft. Danach wird aufgezeigt, welche technischen Implementierungen für den Test erforderlich waren und wie diese umgesetzt wurden. Im Anschluss folgt die Beschreibung der praktischen Experimente, die durchgeführt wurden, um die Effekte der in Kapitel 2 vorgestellten UX-Gesetze und UX-Prinzipien auf die Usability und User Experience zu untersuchen. Jedes dieser Konzepte aus dem Bereich der HCI wird kurz zusammengefasst, um das jeweilige Prinzip in Erinnerung zu rufen. Darauf folgt eine Erläuterung des jeweiligen Versuchsaufbaus und der hierfür eingesetzten Webschnittstellen und Interaktionselemente.

Nach der Beschreibung des jeweiligen Testaufbaus werden die Metriken vorgestellt, anhand derer die Usability und User Experience während des Tests erfasst und im Anschluss bewertet wurden. Daran schliessen sich die Präsentation und Interpretation der Ergebnisse sowie mögliche Erklärungen für die Beobachtungen in den jeweiligen Versuchstests an. Mithilfe der in Kapitel zwei vorgestellten UX-Gesetze, Gestaltungsprinzipien und menschlichen Wahrnehmungsmustern wird versucht, die Ergebnisse zu interpretieren resp. zu begründen.

4.1 Durchführungsmethode des Usability- und User Experience-Tests

Im Rahmen dieser Arbeit wurde ein Usability- und User Experience-Test mit insgesamt 37 Teilnehmern durchgeführt. Die gewählte Testmethode kombiniert Elemente eines klassischen, moderierten Tests mit Aspekten eines unmoderierten Tests. Sowohl moderierte als auch unmoderierte Usability-Tests wurden in Kapitel 3 Abschnitt 3.1 detailliert beschrieben.

Die Durchführung des Usability- und User Experience-Tests erfolgt auf einer eigens dafür entwickelten, voll funktionsfähigen Webseite. Diese ermöglicht es den Teilnehmenden, die vorgesehenen Versuchsexperimente auf der Webseite eigenständig und im Rahmen einer vollständigen und individuellen User Experience zu absolvieren. Um unterschiedliche User Experiences gewährleisten zu können, werden die Teilnehmenden in drei Gruppen aufgeteilt.

Der Test erfolgt in einer Eins-zu-eins-Situation (One-to-one-Session). Die Teilnehmenden werden dabei vom Versuchsleiter direkt beobachtet, äussern jedoch keine laut gedachten Kommentare zu ihrem Vorgehen – die Think-Aloud-Methode wird also nicht angewendet. Während des Tests greift die Testleitung in der Regel nicht ein und stellt keine Fragen, um Ablenkungen zu vermeiden. So sollen sich die Teilnehmenden ganz auf die Nutzung und Interaktion konzentrieren können. Auffälliges Verhalten oder erkennbare Probleme (Usability-Issues) während der Testdurchführung werden vom Versuchsleiter dokumentiert. Die Erhebung der Daten zur späteren Bewertung der Usability, der User Experience sowie des allgemeinen Nutzerverhaltens erfolgt automatisiert im Hintergrund durch das System. Alle relevanten Daten werden dabei automatisch in einer SQL-Datenbank gespeichert.

Nach einigen durchgeführten Versuchsexperimenten wird zusätzlich eine kurze Nutzerbefragung durchgeführt. Diese Nachbefragung betrifft den durchgeführten Versuchstest und verlangt eine subjektive Bewertung unterschiedlicher Aspekte. Dazu beantworten die Teilnehmenden fünf Fragen in einem webbasierten Formular. Die Fragen wurden in leicht angepasster Form dem User Experience Questionnaire (UEQ) entnommen. Was genau bewertet wird, wird in diesem Kapitel unter Abschnitt 4.2 beschrieben. Im Anschluss an den Usability- und User Experience-Test wird eine kurze Nachbefragung der Teilnehmenden durchgeführt, um Einblicke in die Beweggründe und Entscheidungsprozesse ihres Verhaltens zu erhalten und potenzielle Usability-Probleme, die während des Tests beobachtet wurden, nachvollziehen zu können.

Die eben beschriebene Vorgehensweise wurde bewusst gewählt, da es sich bei

dem Test nicht nur um eine isolierte Bewertung der Usability einzelner Webschnittstellen handelt. Stattdessen steht die Analyse einer möglichst realitätsnahen, ganzheitlichen User Experience im Mittelpunkt, die von den Erfahrungen und dem Vorwissen der Nutzer beeinflusst wird.

Zu beachten ist, dass das Alter der Probanden für die Durchführung des Usability- und User Experience-Tests und auch in der späteren Ergebnispräsentation und -interpretation nicht berücksichtigt wird. Etwaige Verzerrungen, die sich aus Unterschieden zwischen den Altersgruppen ergeben könnten, werden somit in Kauf genommen.

4.2 Aufbau und Abfolge des Usability-Tests

Für die Durchführung der Versuchsexperimente im Rahmen des Usability-Tests wurde, wie bereits erwähnt, eine vollständige Webseite entwickelt. Diese verfügt über eine Navigationsstruktur sowie weitere Funktionalitäten, die von den Teilnehmenden individuell genutzt werden können. Die Webseite enthält verschiedene Kategorien von Experimenten, innerhalb derer die jeweiligen Versuche als auswählbare Elemente dargestellt sind. Die Testpersonen sollen die Webseite eigenständig verwenden, um alle Versuchsexperimente in den jeweiligen Kategorien durchzuführen. Eine Experimentreihe gilt dann als abgeschlossen, wenn alle zugehörigen Versuche innerhalb einer Kategorie beendet wurden. Die Webseite dient dabei als zentrale Navigationsstruktur, mit der die Teilnehmenden selbstständig die Kategorien finden, Experimente auswählen und durchführen können. Sie gestalten also durch die Nutzung ihre eigene User Experience.

Ziel dieses Aufbaus ist es, nebst der Untersuchung der UX-Gesetze auch eine möglichst realistische und ganzheitliche User Experience zu simulieren – eine Erfahrung, die durch das Verhalten, die Erfahrungen, Vorlieben und das Vorwissen der Nutzer geprägt ist. Die Teilnehmenden entscheiden selbst, wie sie bei der Durchführung der Experimente vorgehen. Dadurch entsteht im Verlauf des Tests ein individuelles Nutzungsmuster. Die Webseite ermöglicht die eigenständige Durchführung, ohne dass der Versuchsleiter eingreifen muss. Dafür stehen unterschiedliche Hilfestellungen zur Verfügung, wie etwa automatisch generiertes Feedback, das die Teilnehmenden bei Bedarf stets über die nächsten Schritte informiert.

Die 37 Teilnehmenden werden in drei Gruppen – A, B und C – eingeteilt, um verschiedene User Experiences zu erzeugen. Gruppe A enthält 14 Teilnehmer, Gruppe B 15 und Gruppe C 8 Teilnehmer. Je nach Gruppe unterscheidet sich die Darstellung und das Verhalten der Webseite, etwa in Bezug auf Aufbau, Navigation oder

Feedback. Dadurch entstehen unterschiedliche Nutzungsszenarien, die schliesslich miteinander verglichen werden können.

An dieser Stelle soll angedeutet werden, dass es neben den eigentlichen Experimenten auch eine Metaebene des Usability-Tests gibt: Hier wird nicht nur untersucht, wie die Nutzer einzelne Aufgaben bearbeiten, sondern auch, wie sie sich zwischen den Experimenten auf der Webseite bewegen. Diese Metaebene zielt darauf ab, zu analysieren, wie die Teilnehmenden mit der Struktur, Navigation und Logik der Webseite umgehen und wie sie diese Aspekte wahrnehmen, um den gesamten Test abzuschliessen. So liefert der Test nicht nur Einblicke in das Verhalten innerhalb einzelner Experimente, sondern auch in die Art und Weise, wie Nutzer digitale Systeme insgesamt nutzen.

Nutzerbefragung

Im Anschluss an ausgewählte Versuchsdurchläufe, welche die Teilnehmenden im Rahmen der Interaktion absolvieren, erfolgt eine kurze Befragung. In dieser bewerten die Nutzer den Test anhand von fünf Fragen individuell. Diese Fragen stammen in leicht angepasster Form aus dem sogenannten User Experience Questionnaire (UEQ), einem Standardfragebogen zur Messung und Beurteilung von Usability und User Experience [28].

Die fünf Fragen beziehen sich auf zentrale Aspekte der Nutzerwahrnehmung während des Versuchstests: Die kognitive Belastung während der Aufgabenerfüllung, die Erlernbarkeit und Einprägsamkeit der eingesetzten Schnittstelle und Systemfunktionen sowie das mentale Modell der Nutzer und die wahrgenommene Konsistenz der Schnittstelle. Die Antworten zu den ersten beiden Fragen erfolgen auf einer Skala von 1 bis 5 (1: sehr gering, 5: sehr hoch). Die erste Frage zielt darauf ab, die wahrgenommene kognitive Belastung zu erfassen. Die Nutzer sollen beurteilen, wie viel Aufwand, Mühe, Aufmerksamkeit oder Konzentration erforderlich war, um die Aufgabe abzuschliessen.

Die zweite Frage prüft, inwieweit die grafische Benutzeroberfläche dem mentalen Modell des Nutzers entsprach: Funktionierte die Schnittstelle so, wie es der Nutzer erwartete, und konnte er vorhandenes Wissen aus früheren Erfahrungen anwenden, um die Aufgabe zu lösen.

Die dritte Frage bewertet die Verständlichkeit und intuitive Nutzbarkeit der Benutzeroberfläche. Auch hier erfolgt die Antwort auf einer Skala von 1 bis 5, wobei ein hoher Wert für eine besonders einfache und selbsterklärende Bedienbarkeit steht.

Die vierte Frage bezieht sich auf die Erlernbarkeit, Erkennbarkeit und Einprägsamkeit der grafischen Oberfläche. Die Teilnehmenden beurteilen auf einer Skala von 1 (sehr schwierig) bis 5 (sehr einfach), wie leicht es ihnen fiel, die Funk-

tionsweise der Oberfläche zu verstehen, wiederzuerkennen und im Gedächtnis zu behalten.

Die fünfte Frage verlangt eine Einschätzung der Konsistenz. Das heisst, der Nutzer bewertet, ob sich die Schnittstelle intern wie auch extern konsistent verhält. Das Formular mit den beschriebenen Fragen ist im Anhang C unter der Bezeichnung „Umfrageformular“ in Abbildung C.2 abgebildet.

4.2.1 Technische Implementierung der Testwebseite

Die technische Umsetzung der Webseite bildet die Grundlage für die Durchführung und Analyse der Usability- und User Experience-Experimente. In diesem Abschnitt wird erläutert, mit welchen Technologien die Webseite entwickelt wurde, wie die einzelnen Komponenten aufgebaut sind und wie sie miteinander interagieren. Dabei wird das System in Frontend, Backend und Datenbank unterteilt und die Kommunikation zwischen diesen Ebenen näher beschrieben. Eine grafische Illustration der Gliederung in die drei Schichten der Webseite ist in Anhang B unter „Diagramme der Implementierung“ abgebildet.

Grafische Benutzeroberfläche (GUI/Frontend)

Die grafische Benutzeroberfläche der Webseite wurde mit dem Framework Angular¹ umgesetzt. Angular eignet sich für komplexe, interaktive Single-Page-Webanwendungen, da es auf einer komponentenbasierten Architektur basiert. Diese Struktur unterstützt eine klare Trennung von Logik und reiner HTML-Darstellung und erleichtert die Wiederverwendbarkeit von Quellcodemodulen. Die Implementationsstruktur, also die Zusammensetzung der Angular-Komponenten, wird in Anhang B, Abbildung B.2 abgebildet. Die Benutzeroberfläche, mit der die Testpersonen interagieren, wurde vollständig im Frontend realisiert. Die Nutzer navigieren durch die Experimentkategorien, wählen Versuche aus und erhalten während der Durchführung visuelles Feedback. Alle Eingaben und Aktionen im Frontend werden über HTTP-Anfragen an das Backend gesendet, welches nachfolgend beschrieben wird.

Serverlayer (Backend)

Das Backend wurde mit C# und dem .NET Framework² entwickelt. Es bildet die logische Businessschicht der Anwendung und ist verantwortlich für die Bearbeitung der Http-Anfragen aus dem Frontend sowie für die Verwaltung der Nutzer- und Testdaten. Über klar definierte REST-Endpunkte stellt das Backend die notwendigen

¹Webframework von Google: <https://angular.dev/>, abgerufen am: 08.06.2025

²Microsoft .Net, <https://dotnet.microsoft.com/en-us/>, abgerufen am 27.06.2025

Funktionen zur Verfügung, beispielsweise zum Speichern von Testergebnissen oder zum Laden experimentbezogener Inhalte. Die strukturierte Trennung der Anwendungslogik vom Frontend ermöglicht eine bessere Wartbarkeit und Erweiterbarkeit des Systems.

Datenbanklayer

Die Datenbank dient als persistente Speicherung aller Daten. Hierzu zählen unter anderem Nutzerinformationen, durchgeführte Testläufe, Navigationspfade, Metriken und die Gruppenzugehörigkeit der Teilnehmenden. Zum Einsatz kommt eine relationale SQL-Datenbank, die für diese Art strukturierter und verknüpfter Daten besonders geeignet ist. Die Datenbank bildet die Grundlage für die spätere Auswertung der Usability- und User Experience-Metriken sowie für die Analyse der Metaebene des Nutzerverhaltens.

Kommunikation zwischen den Layern

Die Kommunikation zwischen den einzelnen Schichten erfolgt über moderne und etablierte Schnittstellen. Das Frontend kommuniziert über eine REST-API mit dem Backend. Dabei werden die ausgetauschten Daten im JSON-Format über HTTP-Anfragen ausgetauscht. Die Verbindung zwischen Backend und Datenbank wird mithilfe des Entity Framework Core hergestellt, einer objekt-relationalen Abbildungstechnologie (ORM). Diese erlaubt den Zugriff auf Datenbankinhalte über C#-Objekte, was den Entwicklungsaufwand reduziert und für Konsistenz in der Datenverarbeitung sorgt. Durch diese klare Trennung der Schichten kann jede Ebene unabhängig weiterentwickelt, getestet und analysiert werden.

4.2.2 Aufbau und Funktionsweise der Testwebseite

Grundlegende Funktionsweise

Die Webseite enthält ein Navigationsmenü mit Links zu den zentralen Bereichen der Seite: Experimente, Einstellungen und Hilfe. Die angezeigte Navigation passt sich abhängig vom angemeldeten Probanden und dessen Gruppenzugehörigkeit an. Die unterschiedlichen Varianten der Navigation sind in Anhang C unter Navigationsstrukturen dargestellt. Zusätzlich zeigt der Navigationsbereich die E-Mail-Adresse des aktuell angemeldeten Nutzers sowie einen Link zur Abmeldung an. Über den Menüpunkt Experimente gelangt der Nutzer zur Hauptseite der Webseite, auf der eine Übersicht aller verfügbaren Experimentkategorien dargestellt wird. Jede dieser Kategorien repräsentiert eine Experimentreihe mit mehreren zugehörigen Versuchs-

experimenten. Die Liste der Kategorien kann sortiert werden, so dass abgeschlossene zuerst und anschliessend noch nicht abgeschlossene Kategorien angezeigt werden.

Jede Experimentreihe wird auf der Übersichtsseite als rechteckiges, anwählbares Element visualisiert. Diese Elemente informieren die Nutzer darüber, ob die Experimentreihe bereits freigeschaltet wurde, wie viele Versuchstests sie enthält und wie viele davon bereits abgeschlossen wurden. Sobald alle Versuchsexperimente einer Experimentreihe beendet sind, wird dies visuell hervorgehoben: Das entsprechende Rechteck erhält einen grünen Rahmen, und in der rechten oberen Ecke erscheint ein Häkchen-Symbol (vgl. Anhang C, Screenshots der Webseite, Abbildung C.7). Zusätzlich wird der Status der jeweiligen Experimentreihe automatisch auf abgeschlossen gesetzt. Durch diese visuelle Rückmeldung wird den Nutzern auf einfache Art und Weise vermittelt, welche Experimente sie bereits erfolgreich abgeschlossen haben und welche noch zur Bearbeitung ausstehen. Ergänzend zur Darstellung aller Experimentkategorien befindet sich in der rechten unteren Ecke des Bildschirmbereichs ein blauer, runder Button mit einem Informationssymbol. Wird dieser Button betätigt, öffnet sich ein Dialogfenster mit hilfreichen Tipps. Dieses Dialogfenster bietet allgemeine Informationen über den Aufbau und die Inhalte der Webseite sowie über verfügbare Einstellungen und unterstützende Hilferessourcen. Dieser Button kann während des gesamten Interaktionsprozesses betätigt werden.

Über einen weiteren, deutlich sichtbaren blauen Button am unteren Bereich jeder Experimentreihe mit der Aufschrift „Zu den Versuchsexperimenten“ kann der Nutzer zur Übersicht der Versuche innerhalb einer ausgewählten Experimentreihe navigieren. Dort werden alle zugehörigen Versuchsexperimente in Listenform angezeigt (vgl. Anhang C, Versuchsliste, Abbildung C.10). Jedes Listenelement enthält einen Button mit der Bezeichnung „Zu den Versuchsdetails“, über den die Nutzer zu einer ausführlichen Beschreibung des jeweiligen Versuchstests gelangen. Wird anschliessend der grüne Button „Versuch starten“ betätigt, beginnt die Durchführung des entsprechenden Versuchs.

Nach dem Start wird die eigentliche Schnittstelle des Versuches direkt in die Webseite eingebettet. Während des laufenden Versuchs wird im oberen linken Bereich eine Versuchsanleitung konstant angezeigt (vgl. Anhang D Abbildung D.1). Diese kann per Klick geöffnet oder geschlossen werden. Der Nutzer kann also entscheiden, ob und wann er diese konsultieren will. Somit sind diese Instruktionen eine Art modeless Feedback. Die Anleitung führt die Testperson schrittweise durch den Ablauf des Experiments. Je nachdem, welchen Fortschritt der Nutzer innerhalb des Versuchs erzielt, passt sich die Anleitung dynamisch an und aktualisiert die angezeigten Schritte automatisch. Auf diese Weise wird sichergestellt, dass die Nutzer jederzeit über den aktuellen Stand und das weitere Vorgehen informiert

sind. Die dynamische Unterstützung trägt wesentlich dazu bei, eine selbstgesteuerte Durchführung der Versuchsexperimente zu ermöglichen, ohne dass ein Eingreifen durch den Testleiter nötig ist.

Solange ein Nutzer sich in einem aktiven Versuch befindet, unterbindet das System die Navigation zurück zur Hauptwebseite. Dies verhindert, dass ein Nutzer ein aktives Experiment unterbrechen und später wieder fortsetzen kann. Denn dies würde zu verfälschten Daten führen. Das System erkennt, wann der Endzustand eines Experimentes erreicht wurde und beendet in diesem Fall das Experiment automatisch. Wenn eine Nutzerbefragung für den abgeschlossenen Test vorgesehen ist, wird dem Nutzer nach Abschluss des Tests das in Abschnitt 4.2 beschriebene Formular visualisiert. Sobald die Fragen im Formular beantwortet und bestätigt sind, wird der Nutzer zurück zur Versuchsliste der Experimentenreihe zurücknavigiert. Von dort aus kann der angemeldete Nutzer frei entscheiden, wie er mit der Durchführung der Experimente fortfahren will.

Einstellungen

Die Navigationsstruktur der Webseite beinhaltet einen Menüpunkt mit der Bezeichnung „Einstellungen“. Nach dem Aufruf dieses Bereichs gelangen die Webseitennutzer zu den konfigurierbaren Optionen der Webseite. Insgesamt stehen drei Einstellungsmöglichkeiten zur Verfügung: Die erste Option erlaubt es, die progressive Offenlegung der Experimentkategorien zu aktivieren oder zu deaktivieren. Die zweite Einstellung bezieht sich auf die progressive Offenlegung der einzelnen Versuchstests innerhalb einer Experimentkategorie. Die dritte Option ermöglicht es, die automatische Auswahl des jeweils nächsten Versuchstests innerhalb einer ausgewählten Kategorie ein- oder auszuschalten. Die Seite der Einstellungen ist im Anhang C in Abbildung C.9 abgebildet.

Ist die progressive Offenlegung der Experimentkategorien deaktiviert, werden den Nutzenden alle Kategorien zur freien Auswahl angezeigt. In diesem Fall besteht keine Einschränkung hinsichtlich der Reihenfolge, in der die Kategorien bearbeitet werden. Sobald jedoch die progressive Offenlegung aktiviert ist, ändert sich dieses Verhalten: Nun erscheinen nur noch bereits abgeschlossene Kategorien sowie genau eine weitere, noch nicht absolvierte Kategorie, deren Bearbeitung als nächster Schritt vorgesehen ist (vgl. Abbildung D.15 in Anhang D). Der Nutzer wird gezwungen, die ihm präsentierte Kategorie als nächste auszuwählen. Dadurch erfolgt der Fortschritt der Nutzer schrittweise bzw. linear. Ein freier Entscheidungsprozess entfällt somit. Dasselbe Verhalten zeigt das System auch innerhalb der Experimentkategorien, sofern dort die progressive Offenlegung für die enthaltenen Versuchstests aktiviert wurde. Auch hier werden nur die jeweils abgeschlossenen Tests

und der nächste zur Bearbeitung vorgesehene Test angezeigt, während alle übrigen zunächst ausgeblendet bleiben. Diese Konfigurationsmöglichkeiten dienen dazu, das Nutzerverhalten gezielt zu steuern (Benutzerführung) und verschiedene Varianten von Nutzererfahrungen zu untersuchen.

Im nachfolgenden Abschnitt werden die im Rahmen des Usability- und User Experience-Tests durchgeführten Experimente beschrieben sowie deren Ergebnisse präsentiert und interpretiert.

4.3 Evaluierung der Navigationseffizienz

Wie in der Beschreibung zur Durchführung des Usability-Tests erläutert wurde, nutzten die Probanden für die Bearbeitung der Versuchsexperimente eine voll funktionsfähige Webseite. Im Rahmen der Tests mussten die Teilnehmenden wiederholt Navigationsprozesse innerhalb dieser Webseite durchlaufen.

Zunächst musste eine Experimentkategorie ausgewählt werden, daraufhin folgte die Wahl eines spezifischen Versuchstests innerhalb der gewählten Kategorie. Anschliessend sollten die Teilnehmenden die zugehörige Versuchsbeschreibung lesen und danach den Test starten. Nach Abschluss des Tests wurden sie automatisch zur Übersichtsseite der zuvor gewählten Experimentkategorie zurücknavigiert. Von dort aus konnte zurück zur Kategorielliste navigiert werden, um eine andere Kategorie auszuwählen oder der nächste Test der ausgewählten Kategorie gestartet werden. Sobald alle Tests innerhalb einer Kategorie abgeschlossen waren, mussten die Teilnehmenden zur Hauptseite, auf der sämtliche Kategorien aufgelistet sind, navigieren, um einen nächsten Handlungszyklus zu beginnen.

Die Gestaltung der Webseite ermöglicht es, die beschriebenen Navigationsprozesse mit einer (optimalen) Effizienz zu absolvieren – also mit einer minimalen Anzahl notwendiger Mausklicks und Seitenaufrufe. Da das Design der Webseite den Teilnehmenden vor dem Test unbekannt war und sie erstmals im Rahmen des Experiments damit interagierten, ist es von Interesse zu untersuchen, mit welcher Effizienz sie die Navigation durchführten und ob sie dabei ein Verhalten zeigten, das eine möglichst effektive Nutzung der Navigationsstruktur erlaubte.

Die Ergebnisse dieser Untersuchung könnten erkenntnisreich sein, da sie Aufschluss darüber geben können, in welchem Ausmass die Navigationslogik der Webseite erlernt und angewendet wurde.

4.3.1 Usability und User Experience-Metriken

Benutzer-Lostness

Für die Untersuchung der Navigationseffizienz wird der sogenannte Lostness-Faktor L berechnet und analysiert. Dieser beschreibt, wie effizient ein einzelner Navigationszyklus, wie er oben beschrieben wurde, durchgeführt wurde. Der berechnete Wert gibt Aufschluss darüber, ob ein Webseitenutzer aus Effizienz­sicht orientierungslos war oder nicht. Der Lostness-Faktor L ist in [22] detailliert beschrieben und durch die folgende Formel gegeben:

$$L = \sqrt{\left(\frac{N}{S} - 1\right)^2 + \left(\frac{R}{N} - 1\right)^2} \quad (4.1)$$

Dabei gilt:

- N : Anzahl unterschiedlicher aufgerufener Seiten und ausgeführter Mausklicks während eines Navigationsprozesses (doppelt aufgerufene oder angeklickte Schaltflächen werden nicht gezählt)
- S : Totale Anzahl aufgerufener und ausgeführter Klicks während eines Navigationsprozesses (mehrfach aufgerufene Seiten oder Klicks auf Schaltelemente werden gezählt)
- R : Die minimale (optimale) Anzahl von Seiten oder Klicks, die aufgerufen bzw. ausgeführt werden müssen, um einen Navigationszyklus abzuschliessen.

Gemäss P. A. Smith in [29] entspricht ein L -Wert von 0.0 einer optimalen Navigationseffizienz. Werte im Bereich von $0.1 \leq L < 0.2$ weisen auf eine sehr gute Effizienz hin. Liegt der L -Wert zwischen $0.2 \leq L < 0.4$, kann weiterhin von einer guten Effizienz ausgegangen werden, jedoch nicht mehr optimal. Wird ein Wert im Bereich von $0.4 \leq L < 0.5$ erreicht, deutet dies darauf hin, dass der Navigationsprozess nicht mehr effizient durchgeführt wurde. Überschreitet der L -Wert den Wert von 0.5 ($0.5 \leq L$), gilt der Nutzer als stark orientierungslos. Die Navigationsstruktur wurde in diesem Fall nicht wie vorgesehen genutzt.

4.3.2 Ergebnisse

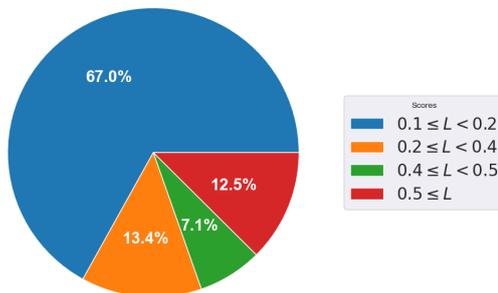
Verteilung der Benutzer-Lostness

Abbildung 4.1 visualisiert links die Verteilung der mittleren Benutzer-Lostness-Werte über alle Teilnehmer hinweg. Rechts dagegen wird veranschaulicht, wie viele

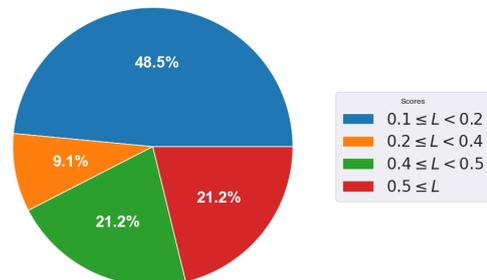
Nutzer einen Lostness-Wert in einem bestimmten maximalen Wertebereich während mindestens eines Navigationsprozesses erreichten. Abbildung 4.1b zeigt somit, wie viel Prozent der Probanden zum Beispiel mindestens einmal einen Lostness-Wert zwischen 0.4 und 0.5 oder mehr als 0.5 erreichten.

Abbildung 4.1 links zeigt, dass etwa 80% der Webseitenutzer einen mittleren Lostness-Wert zwischen 0.1 und 0.4 haben. Das bedeutet, eine Mehrheit der Teilnehmer navigierte während der gesamten User Experience im Schnitt in einer effizienten Art und Weise und war nicht orientierungslos in der Nutzung der Navigation. Etwa 20% der Nutzer dagegen weisen ein ineffizientes Navigationsverhalten auf, mit Werten zwischen 0.4 und 0.5 oder höher als 0.5. Dieselbe Abbildung rechts zeigt einen anderen wichtigen Aspekt. Die Abbildung zeigt, dass 21.2% der Nutzer während mindestens eines Navigationsprozesses einen Lostness-Wert zwischen 0.4 und 0.5 und ebenfalls 21.2% mindestens einmal einen Lostness-Wert höher als 0.5 erreichten. In etwa 42% der Nutzer nutzten die Navigationsstruktur damit während mindestens einer Navigation eher oder sehr ineffizient. Dieses Verhalten könnte auf ein Usability-Issue hinweisen und wird weiter unten nochmals aufgenommen.

Verteilung der mittleren User-Lostness



Verteilung der maximalen User-Lostness-Wertebereiche



(a) Verteilung der mittleren Benutzer-Lostness-Werte

(b) Verteilung der maximal erreichten Lostness-Werte der Teilnehmenden

Abbildung 4.1: Verteilung der mittleren Benutzer-Lostness und Verteilung der maximalen Wertebereiche der Benutzer-Lostness

Benutzer-Lostness nach Navigationsstruktur

Die Teilnehmer wurden wie bereits erwähnt in drei Gruppen aufgeteilt. Jede Gruppe verwendete eine unterschiedliche Navigationsstruktur. Die Navigationsstrukturen, welche implementiert wurden, sind die folgenden: Mega-Drop-Down-Menü, zwei Typen von Seitennavigationen (Off-Canvas-Menüs) und eine horizontale Navigationsleiste (siehe Anhang C, Navigationsstrukturen). Grafik 4.2 zeigt die mittleren Benutzer-Lostness-Werte, unterteilt nach verwendeter Navigationsstruktur. Wie zu

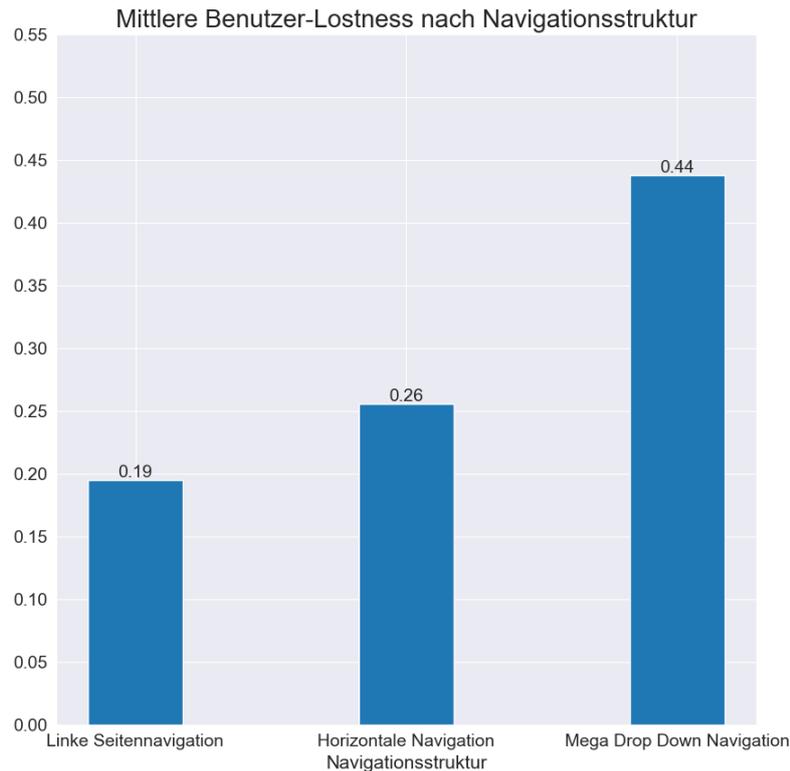


Abbildung 4.2: Durchschnittliche Benutzer-Lostness nach verwendeter Navigationsstruktur

erkennen ist, weisen die Benutzer der linken Seitenmenüs den geringsten mittleren Lostness-Wert auf, gefolgt von den Nutzern der horizontalen Navigationsstruktur. Die zwei Varianten der Seitennavigationsvarianten (Off-Canvas-Menüs) wurden als eine Struktur betrachtet, da die Unterschiede in den beiden Seitennavigationsvarianten zu gering ausfielen. Den höchsten mittleren Lostness-Wert weisen die Nutzer der Mega-Drop-Down-Navigation auf, wie die Grafik zeigt. Die Mega-Drop-Down-Navigation wurde demnach nicht in einer effizienten Art, wie dies möglich gewesen wäre, genutzt.

4.3.3 Interpretation der Ergebnisse

Diese eben präsentierten Ergebnisse sollen nun im folgenden Abschnitt interpretiert und begründet werden.

Usability-Issue

Dass die Nutzer der Mega-Drop-Down-Navigationsstruktur eine hohe Ineffizienz in der Navigation aufweisen, könnte auf ein Usability-Issue hinweisen. Das Mega-Drop-Down-Navigationsmenü erlaubt es, sehr effizient innerhalb der Webseite zu

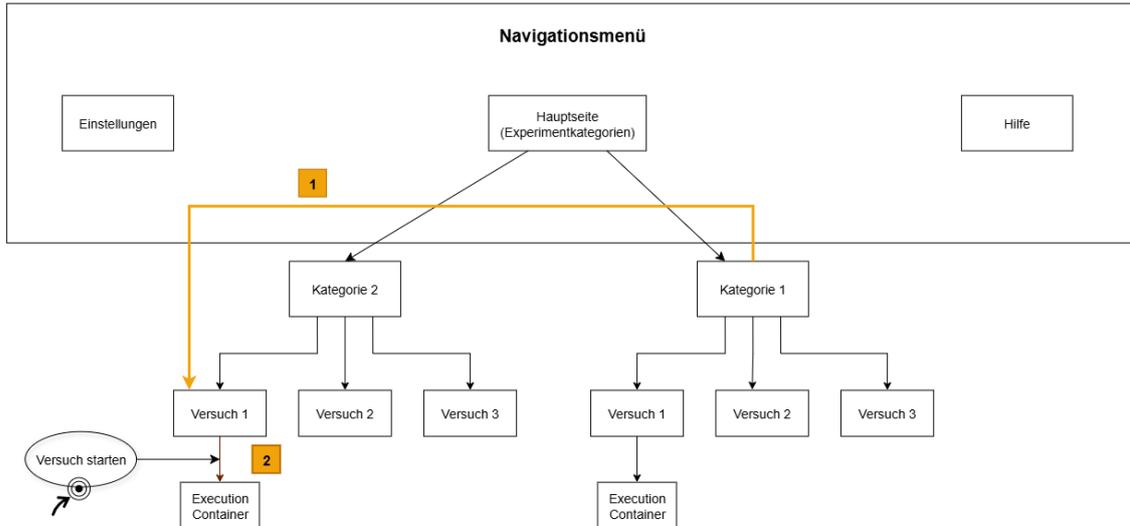


Abbildung 4.3: Optimale Navigation mit dem Mega-Drop-Down-Menü

navigieren. Sofern zum Beispiel alle Versuche einer Kategorie abgeschlossen wurden und ein weiterer Versuch in einer neuen Kategorie gestartet werden will, so müssten hierfür theoretisch maximal zwei Mausklicks ausgeführt und maximal eine Seite aufgerufen werden. Dieser Navigationsprozess ist in Abbildung 4.3 dargestellt. Die braunen Kästchen stellen dabei die Anzahl der Mausklicks dar, um von der einen Kategorie zum nächsten Versuchstest zu gelangen.

Tabelle 4.1 zeigt für jeden Nutzer der Mega-Drop-Down-Navigationsstruktur den Anteil an Navigationsprozessen, bei dem mehr als die doppelte Anzahl der optimalen Anzahl (vier) an Klicks ausgeführt wurden. Wichtig zu bemerken hier ist, dass alle Navigationsprozesse beachtet werden, also auch diejenigen, wo die Kategorie nicht gewechselt wurde.

Teilnehmer	Anteil Navigationen mit 4 oder mehr Mausklicks in %
1	41
2	35
3	59
4	43
5	68
6	36
7	38
8	69
9	45
10	50
11	50

Tabelle 4.1: Anteil der Navigationsprozesse mit 4 oder mehr Mausklicks pro Teilnehmer

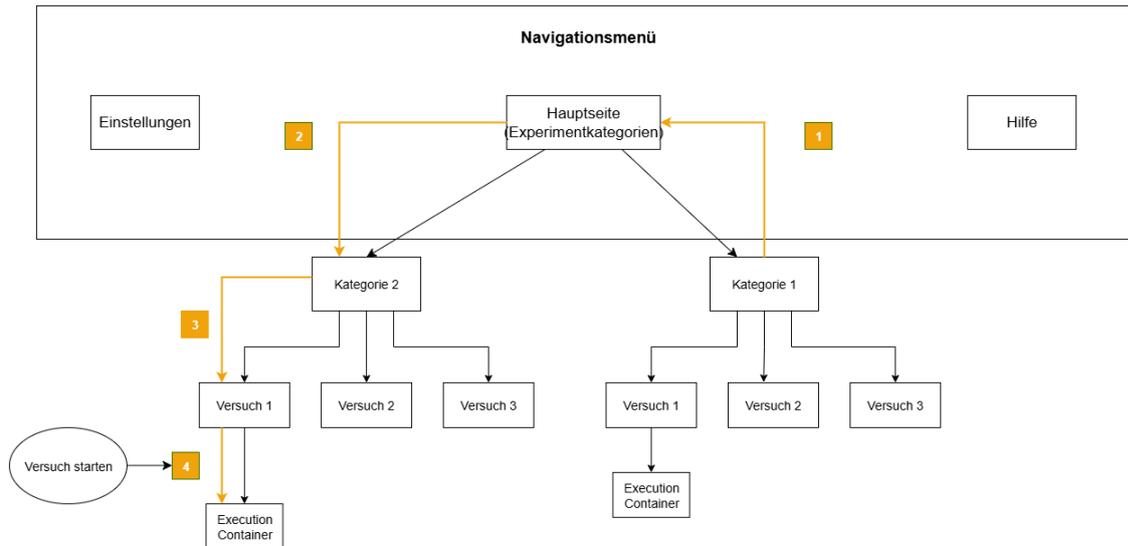


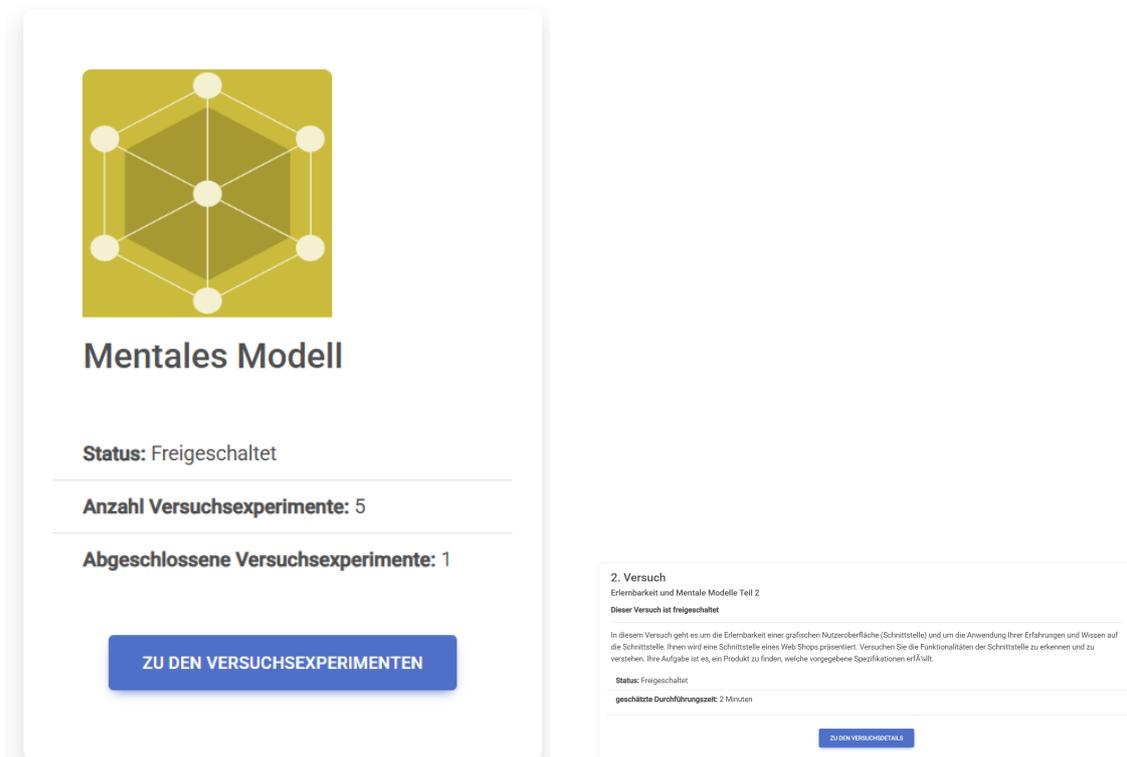
Abbildung 4.4: Gewählte Navigationsmethode der Nutzer der Mega-Drop-Down-Navigationsstruktur

Der Grund für dieses Ergebnis liegt im gewählten Navigationsmuster der Teilnehmer (vgl. Abbildung 4.4), welches während der Interaktion beobachtet wurde: Nachdem die Teilnehmenden eine Experimentreihe abgeschlossen haben, sprangen die Nutzer nicht direkt über das Navigationsmenü zur Detailbeschreibung des nächsten Versuchs innerhalb einer neuen Kategorie, sondern navigierten erst zurück zur Hauptseite. Dort wählten sie dann eine neue Kategorie aus. Anschliessend wurde ein neuer Versuchstest ausgewählt und dann weiter zur Versuchsbeschreibung navigiert. Erst dann wurde der Versuch gestartet. Diese Art von Vorgehen benötigt mehr Mausklicks und Aufrufen von Inhaltsseiten als eigentlich nötig wären und ist somit ineffizienter als das in Abbildung 4.3 dargestellte Vorgehen.

Ein weiteres Usability-Issue, welches während der Tests zeitweise beobachtet wurde, könnte ebenfalls zu den in Abbildung 4.1 dargestellten Ergebnissen, wonach etwa 42% der Probanden in mindestens einem Navigationsprozess als eher oder stark orientierungslos eingestuft wurden, beigetragen haben. Das Problem betrifft das Klickverhalten der Nutzer bei der Auswahl einer Experimentkategorie oder eines Versuchstests. In mehreren Fällen zeigte sich folgendes Verhalten: Um eine Kategorie oder einen Test auszuwählen, muss ein blauer Button am unteren Rand der entsprechenden Box (Kategorie- oder Testbox) betätigt werden (vgl. Abbildung 4.5). Einige Teilnehmer nahmen jedoch die gesamte weisse Box irrtümlicherweise als klickbares Element wahr und klickten sie wiederholt an, ohne dass eine Reaktion durch das System erfolgte. Erst nach mehreren Fehlklicks wurde erkannt, dass die gewünschte Auswahl nur über den Button möglich ist. Dieses Verhalten führte

zu einer erhöhten Anzahl an Mausklicks während des Navigationsprozesses und beeinträchtigte damit die Navigationseffizienz.

Es handelt sich hierbei um ein klassisches Beispiel für eine Diskrepanz zwischen dem mentalen Modell der Nutzer und demjenigen der Entwickler: Das System verhielt sich in diesen Fällen nicht so, wie es aus Sicht der Nutzer zu erwarten gewesen wäre oder wie die Nutzer es aus früheren Erfahrungen kannten.



(a) Eine Experimentkategorie, die zur Auswahl bereit steht

(b) Ein Versuchstest der zur Auswahl bereit steht

Abbildung 4.5: Eine Experimentkategorie oder ein Versuchstest müssen über den blauen Button ausgewählt werden

4.4 Millers Gesetz

In den folgenden Experimenten steht das Gesetz von Miller im Mittelpunkt der Untersuchungen. Dabei werden die beiden Gedächtnisfunktionen Erinnern und Erkennen miteinander verglichen, um ihre Auswirkungen auf die kognitive Belastung und Effizienz während verschiedener Interaktionen zu analysieren. Die kognitive Belastung einer Interaktion wird nach Miller maßgeblich durch diese Gedächtnisprozesse beeinflusst.

Ziel der Untersuchungen ist es, herauszufinden, wie sich das Erinnern und Er-

kennen auf die Usability und User Experience auswirken. Im Fokus steht die Frage, welcher der beiden Prozesse eine schnellere, weniger fehleranfällige und insgesamt effizientere Interaktion ermöglicht. Darüber hinaus soll untersucht werden, ob der Prozess des Erkennens mit einer geringeren kognitiven Belastung einhergeht als das aktive Erinnern. Die Experimente dienen zudem dazu, Unterschiede im Nutzerverhalten bei der Anwendung beider Prozesse systematisch zu erfassen und zu vergleichen.

4.4.1 Testaufbau

Erkennen

Für die Untersuchung des Erkennens wurden zwei Versuchsvarianten entwickelt, die je eine Webschnittstelle eines Web-Shops nutzen. Die grafische Benutzeroberfläche der Web-Shops ist in Anhang D, Gesetz von Miller, Abbildung D.2 dargestellt.

Beide Online-Shops besitzen am linken Rand des Bildschirmbereichs ein textbasiertes Menü mit Links zu den verfügbaren Produktkategorien. Bei Betätigung eines Linkeintrags werden entsprechend Produkte der gewählten Kategorie in der Produktliste rechts vom Menü angezeigt und im Menü werden Links aller Unterkategorien der gewählten Kategorie dargestellt. In der Menükomponente der ersten Versuchsvariante wird dabei die aktuell gewählte Kategorie farblich hervorgehoben und deren Unterkategorien darunter angeordnet. Die Hervorhebung der gewählten Kategorie erfolgt beim Menü der Schnittstelle des zweiten Versuchs bewusst nicht.

Ziel der Teilnehmenden im ersten Versuch ist es, ein Notebook ihrer bevorzugten Marke zu finden, dieses in den Warenkorb zu legen und anschliessend zur Kasse zu navigieren. Für die Produktsuche sollen die Teilnehmenden das zuvor beschriebene textbasierte Menü verwenden. Im ersten Schritt müssen die Teilnehmenden den passenden Menüeintrag identifizieren und auswählen, der sie zur Produktkategorie „IT-und-Multimedia“ führt. Anschliessend gilt es, weitere Unterkategorien zu erkennen und durch deren Auswahl zur entsprechenden Produktkategorie zu gelangen, die „Notebooks“ als Unterkategorie enthält. Sobald die Kategorie „Notebooks“ aufgerufen wurde, wählen die Teilnehmenden ein Gerät ihrer bevorzugten Marke aus der Produktliste aus, legen es in den Warenkorb und begeben sich zur Kasse. Für die Auswahl muss ein Produkt mit der bevorzugte Marke erkannt werden. Der Versuch ist abgeschlossen, sobald das Notebook dem Warenkorb hinzugefügt wurde und der Checkout-Bereich erreicht ist.

Im Zweiten Test müssen die Nutzer eine Gemüsesorte, welche sie für die Zubereitung eines Abendessens benötigen, finden und ebenfalls in den Warenkorb legen. Die Vorgehensweise soll hier diesselbe sein wie im ersten Versuch. Die Teilnehmer

müssen also die Kategorie Lebensmittel erkennen und betätigen, um dann die geeignete Unterkategorie Gemüse zu finden. Schliesslich können die Teilnehmer eine Gemüsesorte auswählen und in den Warenkorb legen und danach zur Kasse gehen. In der zweiten Variante des Tests wird durch die fehlende Hervorhebung der gewählten Kategorie der Erkennungsprozess gezielt erschwert. Denn da die Teilnehmer nun nicht direkt im Menü erkennen, welche Kategorie ausgewählt ist, müssen sie sich allenfalls daran erinnern, welche Kategorie ausgewählt wurde.

Da in diesen Versuchstests das Prinzip Erkennen getestet wird, ist die Suchfunktion in diesen Tests deaktiviert.

Erinnern

Für die Untersuchung der Gedächtnisfunktion Erinnern wurden zwei Versuchstests (zwei Webschnittstellen) entwickelt, in denen die Teilnehmer benötigte Informationen gezielt aus dem Gedächtnis abrufen müssen. Ziel dieser Tests ist die Analyse, inwiefern die Gestaltung von Suchfunktionen, insbesondere das Vorhandensein oder Fehlen von unterstützenden Elementen wie Autovervollständigung - das Nutzerverhalten sowie die Usability beeinflusst.

Im ersten Versuchstest wird den Teilnehmenden die Aufgabe gestellt, ein Smartphone ihrer bevorzugten Marke über das bereitgestellte Suchfeld zu finden. Dieses Suchfeld ist in seiner Funktionalität einfach gehalten: Es bietet keine Autovervollständigung oder Suchvorschläge an. Die Nutzer müssen selbstständig ein Suchwort in das Feld eingeben und über Enter oder über einen Button bestätigen, um ein passendes Ergebnis zu erhalten. Dieser Test fordert von den Teilnehmenden ein hohes Mass an Erinnerungsleistung, da sie sich an den Namen oder die Merkmale der gewünschten Marke eigenständig erinnern und diese korrekt eingeben müssen.

Im zweiten Versuchstest wird eine ähnliche Suchaufgabe gestellt, jedoch unter veränderten Rahmenbedingungen. Die Testpersonen sollen eine Computertastatur finden, die bestimmte, individuell gewünschte technische Spezifikationen aufweist (z.B. periphere Verbindung, Signalübertragung etc.). Anschliessend soll die Tastatur in den Warenkorb gelegt und zur Kasse navigiert werden. In diesem Fall ist die Suchfunktion um eine Autovervollständigungsfunktion erweitert. Während der Eingabe im Suchfeld werden automatisch Produkte, welche die Sucheingaben erfüllen, vorgeschlagen (Siehe Anhang D, Abbildung D.4). Dadurch reduziert sich die Notwendigkeit, Informationen im Gedächtnis zu behalten, da die Benutzer bei der Formulierung ihrer Suchanfrage für sie in Frage kommende Produkte bereits durch die Autovervollständigung erkennen können. Zudem können die vorgeschlagenen Produkte per Mausklick angewählt werden, um direkt zur Detailansicht des gewählten Produktes zu gelangen.

Durch den Vergleich der beiden Versuchstests lässt sich untersuchen, wie stark die Unterstützung durch Funktionen wie Autovervollständigung das Erinnerungsverhalten der Nutzer beeinflusst und in welchem Masse sie zur Effizienz und Zufriedenheit während der Nutzung der Suchfunktion beiträgt.

4.4.2 Usability und User Experience-Metriken

Zur Bewertung verschiedener Aspekte der Usability und User Experience – wie beispielsweise der Anzahl an Fehlern, der Effizienz sowie der kognitiven Belastung – werden für die vier oben beschriebenen Versuchstests verschiedene Performance-Metriken verwendet. Dazu zählen die Auftragszeit und die Anzahl der Fehlklicks sowie subjektive Bewertungen (vgl. Kapitel 3, Abschnitt 3.2.2, Self-reported-Metriken) zur kognitiven Belastung. Auch die Zeit, welche benötigt wird, um einen ersten Klick auf eine Schaltfläche zu tätigen, der den Suchprozess initiiert, wird als Metrik für die Erlernbarkeit eingesetzt.

Darüber hinaus kommen auch vergleichende und kombinierte Metriken zum Einsatz (vgl. Kapitel 3, Abschnitt 3.2.4). Dabei werden insbesondere die gemessenen Auftragszeiten und die subjektiven Bewertungen zur kognitiven Belastung und Erlernbarkeit der Teilnehmenden miteinander in Beziehung gesetzt. Die erfassten Durchführungszeiten werden dabei relativ zur jeweils besten (schnellsten) und schlechtesten (längsten) erreichten Zeit dargestellt und nach Gleichung 3.1 berechnet.

4.4.3 Ergebnisse

Auftragszeit

Tabelle 4.2 gibt einen Überblick über die durchschnittlichen Bearbeitungszeiten sowie die durchschnittliche Anzahl der Fehlklicks in den vier Versuchsdurchläufen über alle Teilnehmenden hinweg. Wie in der Tabelle ersichtlich ist, benötigten die Probanden im ersten Durchlauf der Erkennungsaufgabe mit durchschnittlich circa 48 Sekunden die meiste Zeit. Bereits im zweiten Durchlauf verringerte sich die benötigte Zeit auf etwa 30 Sekunden. Für die beiden Durchläufe, bei denen die Teilnehmenden gezielt ein Suchfeld nutzen mussten (Erinnerungsaufgaben), ergab sich im ersten Versuch eine durchschnittliche Bearbeitungszeit von rund 39 Sekunden, um ein Smartphone der bevorzugten Marke zu finden. Im zweiten Durchlauf stieg die durchschnittliche Ausführungszeit auf etwa 50 Sekunden an.

Insgesamt kann hier also zusammengefasst werden, dass die Erkennungsaufgaben nicht schneller und effizienter als die Erinnerungsaufgaben bearbeitet wurden.

Versuch	Mittlere Auftragszeit in Sekunden	Durchschnittliche Anzahl Fehlklicks
Erkennen 1. Versuch	48.2	1
Erkennen 2. Versuch	30.4	1
Erinnern 1. Versuch	38.8	0
Erinnern 2. Versuch	49.6	1

Tabelle 4.2: Durchschnittliche Anzahl Fehlklicks und Durchführungszeit pro Versuch

Aufwand (Effizienz)

Tabelle 4.9 zeigt die für jeden der vier Versuche die mittlere Anzahl Mausklicks. Wie aus der Tabelle gelesen werden kann, benötigten die Versuchsteilnehmer im ersten Versuch im Schnitt am meisten Mausklicks, um die Aufgabe abzuschliessen. Für die beiden Erinnerungsaufgaben wurden beinahe dieselbe Anzahl an Mausklicks verwendet.

Versuch	Mittlere Anzahl Mausklicks
Erkennen 1. Versuch	9.0
Erkennen 2. Versuch	6.5
Erinnern 1. Versuch	5.2
Erinnern 2. Versuch	5.3

Tabelle 4.3: Mittlere Anzahl Mausklicks je Versuch

Erlernbarkeit

Tabelle 4.4 gibt die durchschnittliche Zeit in Sekunden an, welche die Teilnehmenden benötigten, das erste Mal ein Interaktionselement korrekt auszuwählen – also einen Klick ohne Fehlbedienung auszuführen, der sie im Handlungsablauf einen Schritt weiterführt. Dieser Messwert kann als möglicher Indikator für den Lern- oder Entscheidungsaufwand der Nutzer benutzt werden. Denn diese Zeitmessung misst in einem gewissen Sinne, wie lange die Teilnehmenden brauchten, um Teile der Schnittstellenfunktionalität zu erkennen, zu verstehen und sich für eine bestimmte Vorgehensweise zu entscheiden.

Versuch	Mittlere Zeit bis zum ersten Klick in Sekunden
Erkennen 1. Versuch	16.1
Erkennen 2. Versuch	10.4
Erinnern 1. Versuch	15.8
Erinnern 2. Versuch	16.0

Tabelle 4.4: Mittlere Zeit bis zum ersten Klick auf ein Interaktionselement pro Versuch

Wie der Tabelle zu entnehmen ist, benötigten die Probanden im ersten Erkennungsversuch die meiste Zeit, um eine Entscheidung darüber zu treffen, welches Interaktionselement sie als erstes auswählen sollten, um den Suchprozess voranzubringen. Im zweiten Erkennungsversuch reduzierte sich diese Entscheidungszeit bereits um fast 6 Sekunden und fällt am geringsten aus. In den beiden Erinnerungsaufgaben dagegen fällt die Entscheidungszeit stabil aus, durchschnittlich benötigten die Probanden etwa 16 Sekunden, um eine Entscheidung zu treffen.

Kombinierte und vergleichende Metriken

Wie Tabelle 4.5 zeigt, tritt die größte Diskrepanz zwischen der besten (d.h. schnellsten) und der schlechtesten (d.h. langsamsten) Auftragszeit beim ersten Erkennungsversuch auf. Im Folgenden werden die restlichen Zeitmessungen dieses Versuchs in Relation zu diesen beiden Extremwerten gesetzt. Ergänzend werden die subjektiven Bewertungen der Teilnehmenden hinsichtlich der wahrgenommenen kognitiven Belastung sowie der Erlernbarkeit der grafischen Oberfläche berücksichtigt.

Versuch	Schnellste Zeit in Sekunden	Längste Zeit in Sekunden
Erkennen 1. Versuch	13.2	120.7
Erkennen 2. Versuch	15.6	78.5
Erinnern 1. Versuch	14.8	94.9
Erinnern 2. Versuch	24.0	111.9

Tabelle 4.5: Schnellste und langsamste Ausführungszeit je Versuch

Tabelle 4.6 stellt für 32 Probanden für die Durchführungszeit, die Bewertung der kognitiven Belastung sowie für die Einschätzung der Erlernbarkeit jeweils einen prozentualen Wert dar. Diese Werte wurden gemäss Gleichung 3.1 berechnet. Die Berechnungen von fünf restlichen Probanden konnten nicht sinngemäss gesammelt werden und wurden deshalb weggelassen. Diese Prozentangaben beziehen sich darauf, wie nah der jeweilige Wert am besten erreichten Wert liegt. Ein höherer Prozentwert bedeutet somit mehr Nähe zum besten Wert, der gemessen wurde. Die Tabelle legt nahe, dass teilweise eine Korrelation zwischen der Auftragszeit und den subjektiven Einschätzungen der kognitiven Belastung sowie der Erlernbarkeit der Benutzeroberfläche besteht. Teilnehmende, die mehr Zeit für die Bearbeitung der Aufgabe benötigten, bewerteten diese tendenziell als kognitiv belastender ein und bewerteten die Erlernbarkeit der Oberfläche geringer als diejenigen, die die Aufgabe schneller abgeschlossen hatten.

Ein auffälliges Muster besteht darin, dass Probanden, deren Durchführungszeit unter 75% des Bestwertes liegt, auch bei der Bewertung der kognitiven Belastung und der Erlernbarkeit fast durchgängig Werte von 80% oder weniger aufweisen.

Teilnehmer	Durchführungszeit (%)	Bewertung Kognitive Belastung (%)	Bewertung Erlernbarkeit (%)
1	80.0	100.0	80.0
2	70.0	100.0	100.0
3	70.0	100.0	100.0
4	83.0	80.0	80.0
5	33.0	60.0	80.0
6	66.0	80.0	80.0
7	64.0	80.0	100.0
8	56.0	80.0	80.0
9	64.0	80.0	100.0
10	43.0	60.0	60.0
11	83.0	80.0	60.0
12	79.0	100.0	100.0
13	93.0	60.0	80.0
14	71.0	100.0	80.0
15	70.0	80.0	80.0
16	68.0	80.0	80.0
17	74.0	100.0	80.0
18	35.0	80.0	80.0
19	25.0	100.0	100.0
20	63.0	80.0	80.0
21	67.0	80.0	80.0
22	66.0	80.0	80.0
23	73.0	100.0	100.0
24	0.0	20.0	60.0
25	73.0	40.0	40.0
26	85.0	80.0	60.0
27	90.0	100.0	80.0
28	100.0	80.0	80.0
29	56.0	80.0	100.0
30	59.0	60.0	80.0
31	84.0	80.0	80.0
32	87.0	80.0	80.0

Tabelle 4.6: Bearbeitungszeiten und Bewertungen der kognitiven Belastung und Erlernbarkeit in Prozentwerten

Nutzerbewertungen

Abbildung 4.6 zeigt die durchschnittlichen Nutzerbewertungen zur kognitiven Belastung. Die ersten beiden Erkennungsaufgaben wurden ähnlich bewertet und mit einem durchschnittlichen Score von etwa 2 als eher gering kognitiv belastend eingeschätzt. Die beiden Erinnerungsaufgaben erhielten mit rund 1,8 Punkten eine etwas niedrigere Bewertung und wurden damit als weniger kognitiv belastend empfunden als die ersten beiden Erkennungsaufgaben.

4.4.4 Interpretation der Ergebnisse

Die in Abschnitt 4.4.3 dargestellten Ergebnisse sollen im Folgenden kurz interpretiert werden.

Wie Tabelle 4.2 zu entnehmen ist, weist der erste Erkennungsversuch die höchste durchschnittliche Auftragszeit auf. Diese liegt sogar deutlich über derjenigen des ersten Erinnerungsversuchs, obwohl die Aufgabenkomplexität in beiden Fällen dieselbe war: Im ersten Fall sollten die Probanden ein Notebook ihrer bevorzugten Marke finden, im zweiten Fall ein Smartphone, ebenfalls von ihrer bevorzugten Marke.

Mittlere Bewertungen der kognitiven Belastung nach Versuchstest

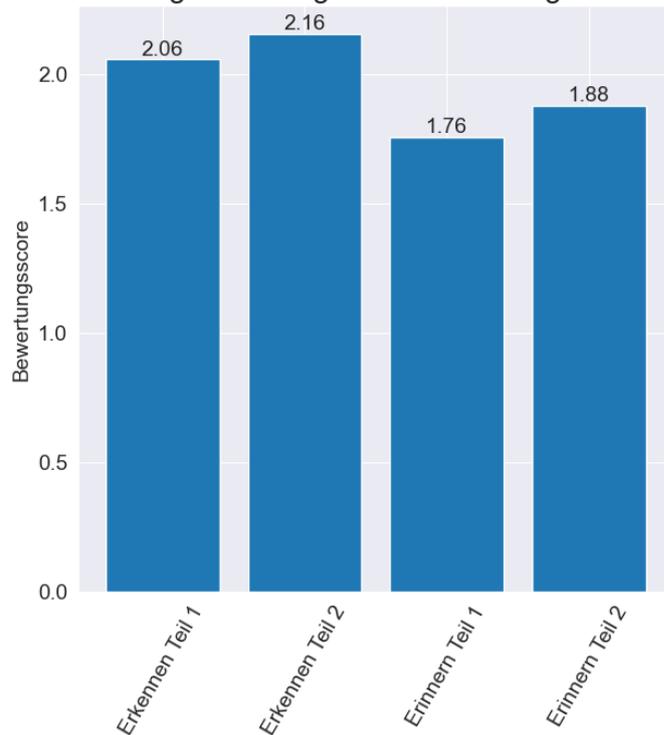


Abbildung 4.6: Mittlere Bewertungen zur kognitiven Belastung der vier Versuchsexperimenten zu Miller's Gesetz

Hierbei ist zu berücksichtigen, dass für viele Teilnehmende die Experimentreihe „Erinnern vs. Erkennen“ den ersten Teil der gesamten User Experience-Studie darstellte. Die vier Versuche dieser Reihe wurden in den meisten Fällen in der in Tabelle 4.2 dargestellten Reihenfolge absolviert. Da die Webshop-Oberfläche in allen vier Aufgaben grundsätzlich identisch aufgebaut war, liegt die Vermutung nahe, dass im ersten Erkennungsversuch zunächst Zeit für das Erlernen der Schnittstellenfunktionalität sowie für das Verstehen der Navigationsstruktur aufgewendet wurde. Diese Annahme wird durch die Ergebnisse in Tabelle 4.4 zusätzlich gestützt. Sie zeigt die durchschnittlichen Sekunden an, welche die Nutzer benötigten, bis sie einen Suchprozess durch einen ersten Klick starteten. Es ist daher plausibel, dass sich zwischen dem ersten und dem zweiten Versuch ein Lernprozess vollzog, der es den Nutzern ermöglichte, die Aufgaben im weiteren Verlauf effizienter und schneller zu bearbeiten (vgl. auch Tabelle 4.3). Dieser Lernprozess trug dann vermutlich auch dazu bei, dass im ersten Erinnerungsversuch die Auftragszeit im Vergleich zum zweiten Erkennungsversuch nicht wieder stark anstieg.

Die Menüschnittstelle der grafischen Oberfläche des zweiten Erkennungsversuchs wurde im Vergleich zur Nutzeroberfläche des ersten Erkennungsversuchs bewusst unterschiedlich gestaltet. In der ersten Version wird die aktuell ausgewählte Ka-

tegorie farblich hervorgehoben, in der zweiten ist dies nicht der Fall. Theoretisch sollte dies die kognitive Belastung erhöhen, da sich Nutzerinnen und Nutzer in diesem Fall eher merken müssen, welche Kategorie ausgewählt ist oder sie müssen diese Information an einer anderen Stelle in der Schnittstelle suchen und erkennen.

Tabelle 4.2 zeigt, dass sich die unterschiedliche Gestaltung kaum negativ auf die Interaktionsgeschwindigkeit der Nutzer auswirkte. Mögliche und plausible Erklärungen könnten einerseits der bereits beschriebene Lernprozess, der weiter oben erwähnt wurde, und andererseits die Fähigkeit der selektiven Aufmerksamkeit, welche ebenfalls zu diesem Resultat beigetragen haben könnte, sein. Während des zweiten Erkennungsversuches mussten die Probandinnen und Probanden insgesamt zwei Links identifizieren. Als erstes den Link „Lebensmittel“ danach den Link „Gemüse“. Sobald der Link „Lebensmittel“ identifiziert und ausgewählt wurde, wurde vermutlich der Fokus direkt auf das Finden und Anklicken des zweiten Links gesetzt, ohne zuvor das Feedback des Systems (farbliche Hervorhebung des Links) abzuwarten.

Ein Usability-Issue, das möglicherweise zum in Tabelle 4.2 dargestellten Ergebnis beigetragen hat, wurde ebenfalls während des ersten Versuchs beobachtet: Einige Probandinnen und Probanden versuchten, das Suchfeld zu verwenden, um das gewünschte Notebook zu finden. Da es sich jedoch um eine Erkennungsaufgabe handelte, war die Suchfunktion bewusst deaktiviert. Das Anklicken der Suchleiste sowie die anschließende Erkenntnis, dass diese Funktion nicht verfügbar war, führten in den meisten Fällen zu einem zusätzlichen Zeitaufwand von mehreren Sekunden und zu mehr Klicks. Ein möglicher Grund für dieses Verhalten könnte darin liegen, dass einige Teilnehmende das Eintippen der bevorzugten Marke und das anschließende Auswählen eines Produkts als effizientere Vorgehensweise empfanden. Dies deutet bereits auf eine Charakteristik mentaler Modelle von Nutzern hin: Bei geringer Aufgabenkomplexität wird die Nutzung einer Suchfunktion offenbar als angenehme und effektive Vorgehensweise wahrgenommen.

Ein Vergleich des zweiten Erkennungsversuchs mit dem ersten Erinnerungsversuch zeigt zudem, dass der Erinnerungsprozess offenbar etwas mehr Zeit beansprucht als der Wiedererkennungsprozess. Vorab sollte hier bemerkt werden, dass die Aufgabenkomplexität dieser beiden Versuche in etwa gleich ist. Das bedeutet, die Anzahl nötiger Ressourcen, um diese Aufgaben abzuschliessen, sind bei beiden Versuchen gleich. Ein möglicher Grund, weshalb der zweite Erkennungsversuch im Schnitt dennoch schneller als der erste Erinnerungsversuch bearbeitet wurde, liegt vermutlich in der unterschiedlichen Aufgabenstruktur und an der Anzahl an mentalen Ressourcen. Beim Erinnern mussten die Teilnehmenden zunächst die Suchfunktion lokalisieren, sich an ihre bevorzugte Smartphone-Marke erinnern, die Suchfunktion anwählen, einen entsprechenden Suchbegriff eingeben und die Suche

auslösen. Deshalb stieg vermutlich auch die Zeit bis zum ersten Klick im Vergleich zu den ersten beiden Erkennungsaufgaben wieder etwas an. Im Gegensatz dazu bestand die Aufgabe beim Erkennen darin, die Kategorie Lebensmittel zu finden und anzuwählen. Danach musste lediglich noch die Kategorie Gemüse im Menü identifiziert werden und anschliessend eine entsprechende Gemüsesorte ausgewählt werden. Insgesamt mussten also zwei Produktlinks lokalisiert werden. Sofern die Funktionsweise des Menüs einmal erlernt wurde, kann die Lokalisierung und Auswahl eines Links effizient und mit wenig mentalem Aufwand erfolgen.

Der in Tabelle 4.6 dargestellte Trend lässt sich möglicherweise dadurch erklären, dass Aufgaben, die über einen längeren Zeitraum hinweg Konzentration und Aufmerksamkeit erfordern, grundsätzlich als kognitiv belastender wahrgenommen werden als solche, die nur eine kurze Aufmerksamkeitsspanne beanspruchen. Das ist eine direkte Konsequenz aus dem Gesetz von Miller. Je höher die empfundene kognitive Belastung einer Aufgabe ist, desto schwieriger erscheint vermutlich auch das Verständnis der notwendigen Schritte zur erfolgreichen Aufgabenerledigung. Dies könnte erklären, weshalb die Erlernbarkeit der Schnittstelle tendenziell geringer bewertet wurde, wenn die Bearbeitungszeit länger ausfiel. Abbildung 4.6 unterstützt zudem die Annahme, dass ein vorgelagerter Lernprozess sowie ein anschliessender Interaktionsprozess, wie sie im ersten Erkennungsversuch bei einigen Teilnehmenden erforderlich waren, zu einer höheren kognitiven Belastung führen können, als wenn direkt mit der Interaktion zur Zielerreichung begonnen werden kann.

4.5 Gesetz von Hick

Das Gesetz von Hick besagt, dass die Zeit, welche benötigt wird, eine Entscheidung zu treffen, mit der Anzahl an Auswahlmöglichkeiten und deren Komplexität steigt (vgl. Kapitel 2, Abschnitt 2.3.1, Gesetz von Hick). Um diesen Aspekt zu untersuchen, wurden drei Experimente entwickelt und durchgeführt.

4.5.1 Testaufbau

Für die drei Experimente wurden drei Schnittstellen eines Online-Webshops implementiert, welche im Aufbau und in der Struktur alle gleich sind (vgl. Anhang D, Gesetz von Hick, Abbildung D.6). Für zwei Komponenten der drei Testschnittstellen variiert dabei die Anzahl an möglichen Auswahlobjekten. Die erste Komponente ist das textbasierte Inhaltsmenü, welches die jeweiligen Produktkategorien als auswählbare Links enthält. Die zweite Komponente ist die tatsächliche Produktliste. In beiden Komponenten wird die Anzahl an Auswahlobjekten dabei über alle

drei Tests schrittweise erhöht. Das heisst, im ersten Experiment werden dem Nutzer in beiden Komponenten am wenigsten Objekte zur Auswahl präsentiert. Im zweiten wird die Anzahl an Produkten um 22 und die Menükomponente um vier Optionen erhöht. Im dritten Experiment werden am meisten Auswahlobjekte präsentiert, 50 Produkte und zwölf Menüeinträge. Die unterschiedlichen Anzahlen an Auswahlobjekten der zwei Komponenten sind in 4.7 abgebildet.

Mithilfe dieser drei Versuchsvarianten soll analysiert werden, wie sich die Entscheidungszeit, welche in allen drei Fällen als eine Art Reaktionszeit interpretiert werden kann, abhängig von der Anzahl an Auswahloptionen verändert. In allen drei Versuchen ist das Ziel der Teilnehmenden dasselbe:

Versuchsnr.	Anzahl Auswahloptionen Menükomponente	Anzahl Produkte
1	4	25
2	8	37
3	12	50

Tabelle 4.7: Anzahl Menüeinträge und Produkte in den drei Versuchen zu Hicks Gesetz

Die Nutzer sind gefragt, insgesamt drei Lebensmittelprodukte auszuwählen, in den Warenkorb zu legen und danach zur Kasse zu gehen. Im ersten Versuch müssen die Probanden drei Lebensmittel für die Zubereitung eines Mittagessens auswählen. Während des zweiten Versuchs entscheiden sich die Probanden für drei Lebensmittel, welche sie für die Zubereitung eines Abendmahls benötigen, und im dritten sollten die Teilnehmer ein Picknick aus drei Lebensmitteln zusammenstellen. In allen drei Versuchen erfolgt die Vorgehensweise auf dieselbe Art und Weise: Als erstes müssen die Probanden im Produktmenü den Eintrag „Lebensmittel“ finden. Dann müssen sie diesen Link anwählen, um alle Lebensmittel angezeigt zu bekommen. Danach erfolgt die Auswahl eines Produktes, welches dann in den Warenkorb gelegt wird. Danach müssen die Teilnehmer wieder über den Link „Lebensmittel“ zurück zur Produktliste navigieren. Dieses Vorgehen wird wiederholt, bis die Teilnehmer drei Produkte im Warenkorb haben. Danach können die Probanden zur Kasse, um den Versuch zu beenden.

4.5.2 Usability und User Experience-Metriken

In den Experimenten zu Hicks Gesetz spielen zeitbezogene Metriken eine zentrale Rolle. Um die Auswirkungen unterschiedlicher Anzahlen von Auswahloptionen auf die Reaktionszeit und die Effizienz der Teilnehmenden zu analysieren, werden verschiedene Metriken verwendet. Zum einen wird die Veränderung der gesamten Auftragszeit betrachtet. Darüber hinaus wird die Zeitspanne, die benötigt wird, um

den Link „Lebensmittel“ das erste Mal zu lokalisieren und auszuwählen, gemessen. Diese Messgrösse dient nicht nur der Erfassung der Reaktionszeit, sondern kann zugleich Hinweise auf die Erlernbarkeit der Benutzeroberfläche liefern. Ergänzend wird die Zeit erfasst, welche die Teilnehmenden für die Auswahl der Produkte aufwenden.

Neben diesen zeitbezogenen Kennwerten werden auch Effizienzmetriken wie die Anzahl der Klicks und Fehlklicks berücksichtigt, die Rückschlüsse auf die Effektivität und Präzision der Nutzerinteraktion unter unterschiedlich vielen Auswahloptionen erlauben. Auch subjektive Bewertungen (Self-reported-Metriken) zur kognitiven Belastung werden untersucht.

4.5.3 Ergebnisse

Auftragszeit

Tabelle 4.8 zeigt die mittleren Bearbeitungszeiten in Sekunden für die drei Versuche zu Hicks Gesetz über alle Teilnehmer hinweg. Wie aus der Tabelle zu lesen ist, benötigten die Teilnehmer die meiste Zeit im ersten Versuch, beinahe 52 Sekunden, um sich für drei Lebensmittel zu entscheiden. Im zweiten Versuch verringerte sich diese Zeit auf rund 42 Sekunden und im dritten und letzten Versuch stieg die Bearbeitungszeit dagegen im Vergleich zum zweiten Versuch wieder leicht um etwa vier Sekunden an.

Versuchsnr.	Durchschnittliche Ausführungszeit in Sekunden
1	51.9
2	42.1
3	46.3

Tabelle 4.8: Durchschnittliche Bearbeitungszeit der drei Versuche zu Hicks Gesetz

Reaktionszeiten und Entscheidungszeiten

Abbildung 4.7 zeigt die mittleren Reaktionszeiten der Teilnehmenden – also wie lange sie benötigten, um nach Beginn des Versuchs den Menüeintrag „Lebensmittel“ zu finden und anzuklicken. Gemessen wurde dies bei Menüs mit vier, acht und zwölf Einträgen. Im ersten und dritten Versuch, also bei vier bzw. zwölf Einträgen, lagen die Reaktionszeiten mit rund neun Sekunden auf einem gleichen Niveau. Im zweiten Versuch mit acht Einträgen war die Reaktionszeit etwas kürzer und lag bei durchschnittlich 6,3 Sekunden.

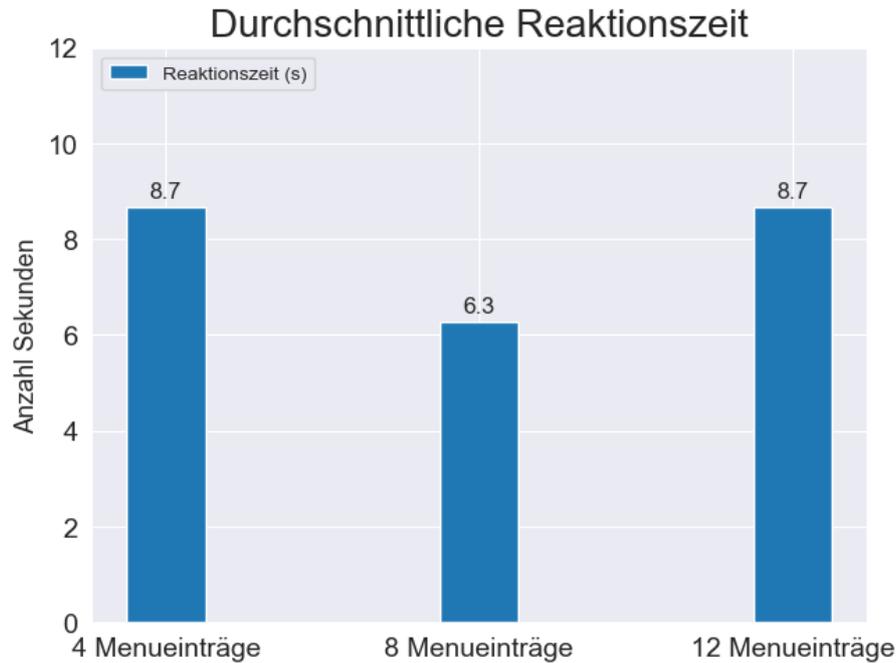
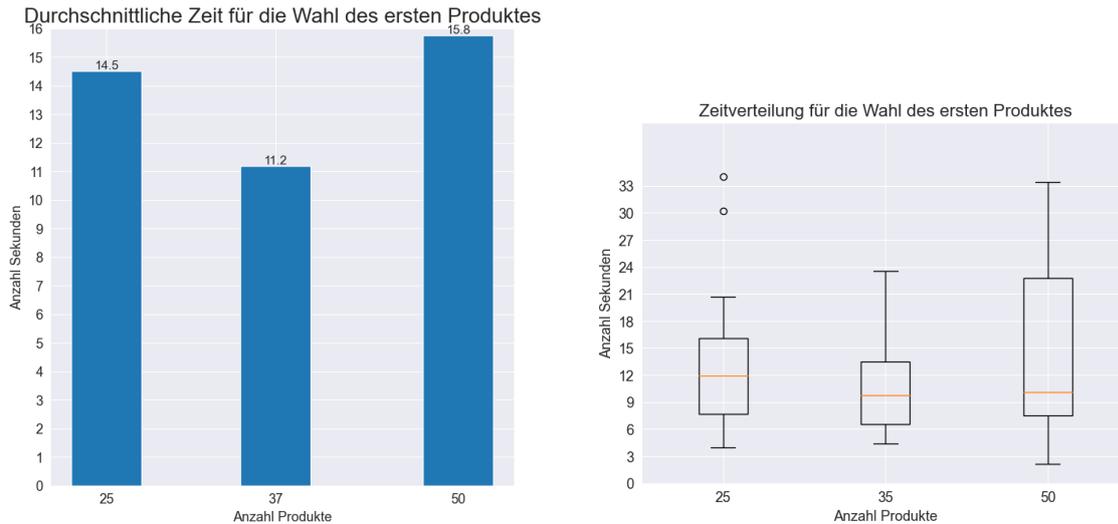


Abbildung 4.7: Reaktionszeiten in Abhängigkeit der Anzahl Menüeinträge

Abbildung 4.8a zeigt die durchschnittlichen Entscheidungszeiten für die Auswahl des jeweils ersten Produkts in den drei Versuchsdurchläufen mit 25, 37 und 50 Produkten. Dabei ist zu erkennen, dass die Teilnehmenden bei 37 Produkten am wenigsten Zeit benötigten, um sich für das erste Produkt zu entscheiden – im Durchschnitt knapp elf Sekunden. Im Vergleich dazu lagen die Entscheidungszeiten bei den beiden anderen Versuchsvarianten bei 14,8 bzw. 14,1 Sekunden, also rund 3,3 bzw. vier Sekunden höher.

Abbildung 4.8b veranschaulicht die Streuung der Entscheidungszeiten über die drei Versuche hinweg. Es zeigt sich, dass die Wertebereiche in den ersten beiden Durchläufen mit 25 bzw. 37 Produkten eine geringere Varianz aufweisen als im letzten Versuch mit 50 Produkten. Die braune Linie innerhalb jeder Box repräsentiert den Median. Auffällig ist, dass der Median in den ersten beiden Versuchen relativ zentral in der Box liegt, was auf eine symmetrische Verteilung der mittleren 50% der Daten und somit auf eine gleichmässige Streuung der Entscheidungszeiten hindeutet. Dies könnte ebenfalls darauf hinweisen, dass es in diesen beiden Durchläufen vergleichsweise wenige Teilnehmende gab, deren Entscheidungszeiten deutlich von denen der übrigen Teilnehmenden abwichen.

Im dritten und letzten Versuch hingegen fällt der Wertebereich deutlich breiter aus. Dies deutet darauf hin, dass einige (ev. wenige) Teilnehmende im Vergleich zu den ersten beiden Versuchen im dritten Versuch wesentlich mehr Zeit benötigten, um sich für das erste Produkt zu entscheiden. Da der Median jedoch in der dritten



(a) Durchschnittliche Entscheidungszeiten für das erste Produkt

(b) Werteverteilung der Entscheidungszeiten für die erste Produktwahl

Abbildung 4.8: Entscheidungszeiten für die Wahl des ersten Produktes

Box relativ weit im unteren Bereich liegt – ähnlich wie im zweiten Versuch – lässt sich vermutlich schlussfolgern, dass der Grossteil der Teilnehmenden vergleichbare Entscheidungszeiten wie in den vorherigen Versuchen aufwies. Die grössere Streuung resultiert vermutlich aus einzelnen Ausreißern mit deutlich längeren Entscheidungszeiten.

Effizienz

Tabelle 4.9 zeigt die durchschnittliche Anzahl an Mausklicks der Teilnehmenden in allen drei Versuchsdurchläufen an. Es ist erkennbar, dass die durchschnittliche Klickanzahl über alle Versuche hinweg nahezu konstant blieb. Die steigende Anzahl an Auswahloptionen in der Menükomponente sowie in der Produktliste führte demnach nicht zu einer Zunahme der benötigten Klicks. Der Interaktionsaufwand für das Abschliessen der Aufgabe blieb somit über alle drei Versuche hinweg nahezu unverändert.

Versuchsnummer	Mittlere Anzahl Mausklicks
1	12
2	11
3	13

Tabelle 4.9: Durchschnittliche Anzahl an Mausklicks während der drei Versuche zu Hicks Gesetz

Nutzerbewertungen

Abbildung 4.9 zeigt die mittleren Bewertungen der kognitiven Belastung der drei Versuchsvarianten zu Hicks Gesetz. Je höher der Wert, desto höher wurde die kognitive Belastung empfunden. Wie die Abbildung zeigt, fallen die Bewertungen zur kognitiven Belastung bei allen drei Tests ähnlich aus. Der letzte Versuch weist sogar die beste Bewertung von 1.5 auf. Die ersten beiden haben praktisch dieselbe Bewertung mit 1.8 bzw. 1.7. Insgesamt wurde also die kognitive Belastung bei allen drei Versuchen als gering bis sehr gering empfunden.

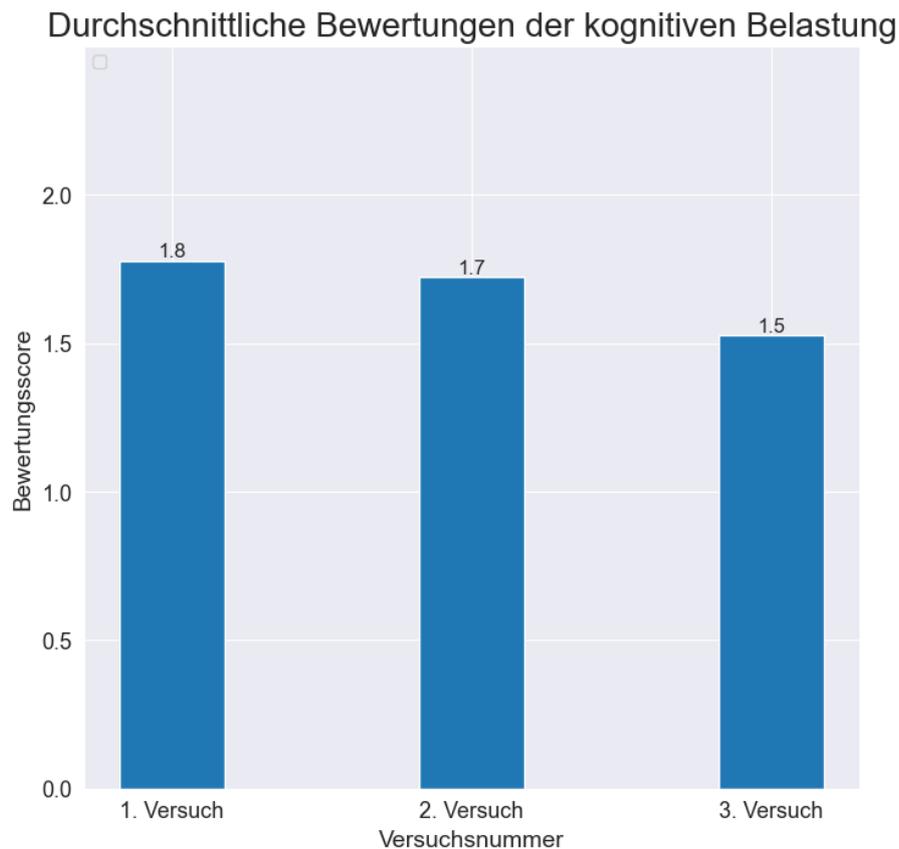


Abbildung 4.9: Durchschnittliche Bewertung der kognitiven Belastung in den drei Versuchen zu Hicks Gesetz

4.5.4 Interpretation der Ergebnisse

Die aus den Experimenten zu Hicks Gesetz gewonnenen Ergebnisse zeigen teilweise unerwartete Ergebnisse, die im Folgenden näher interpretiert und begründet werden sollen. Ein Blick auf Tabelle 4.8 sowie Abbildung 4.7 verdeutlicht, dass eine schrittweise Erhöhung der Anzahl von Produkten und Menüeinträgen nicht zwangsläufig zu einer Verlängerung der Ausführungs- und Reaktionszeiten führt, die Usability

also nicht zwingend negativ beeinflusst. Eine mögliche Erklärung für die stabilen oder im Übergang vom ersten zum zweiten Versuch sogar leicht sinkenden Reaktionszeiten liegt vermutlich in der Kombination von selektiver Aufmerksamkeit und menschlichem Erinnerungsvermögen, resp. der Einprägsamkeit und Erkennbarkeit textbasierter Menüs.

Wie bereits in Kapitel 2, Abschnitt 2.1, Menschliche Faktoren erläutert wurde, ermöglicht die selektive Aufmerksamkeit den Probanden, irrelevante Informationen – in diesem Fall nicht zielführende Links – weitestgehend auszublenden und sich gezielt auf den relevanten, zielführenden Link „Lebensmittel“ zu konzentrieren. Da sich die Position des Ziellinks über alle drei Versuchsvarianten hinweg nicht änderte und die Liste der Links im Menü lediglich durch zusätzliche Einträge am unteren Ende der Liste ergänzt wurde, liegt die Annahme nahe, dass sich die Teilnehmenden die Position des zielführenden Links zusätzlich einprägten und ihn daher im zweiten und dritten Versuch ähnlich schnell erkannten wie im ersten.

Diese beiden Aspekte könnten erklären, weshalb die Reaktionszeiten über die drei Versuchsreihen hinweg weitestgehend unverändert blieben. Ebenso lässt sich damit die konstante Effizienz (Anzahl Mausklicks) der Nutzer begründen, wie Tabelle 4.9 belegt.

Auch Abbildung 4.9 unterstützt diese Interpretation: Die subjektiv bewertete kognitive Belastung steigt mit wachsender Anzahl an Links und Produkten nicht wie erwartet an. Offenbar ermöglichten die selektive Aufmerksamkeit und die Einprägsamkeit des textbasierten, zielführenden Links den Probanden, die Zunahme an Informationsmenge auszublenden, sodass diese nicht als zusätzliche kognitive Belastung empfunden wurde. Die Beibehaltung oder sogar leichte Verbesserung der Interaktionsgeschwindigkeit und Effizienz über die drei Tests hinweg – wie auch in Tabelle 4.8 ersichtlich – dürfte somit ebenfalls auf diese Mechanismen (Einprägsamkeit und selektive Aufmerksamkeit) zurückzuführen sein. Die selektive Aufmerksamkeit scheint ein sehr wirksames Instrument zu sein, die kognitive Belastung trotz der Existenz steigender nicht zielführender Informationen zu reduzieren und sich auf die tatsächlichen Ziele zu fokussieren.

Trotzdem ist nicht auszuschliessen, dass eine wachsende Anzahl von Auswahloptionen die Usability und mitsamt das Interaktionsverhalten einzelner Teilnehmender negativ beeinflusst. So zeigt sich insbesondere in der dritten Versuchsreihe sowohl ein Anstieg der Reaktionszeit (vgl. Abbildung 4.7) als auch der Entscheidungszeit (vgl. Abbildung 4.8a und Abbildung 4.8b) im Vergleich zum zweiten Versuch. Abbildung 4.8b illustriert dabei durch die dritte Box, dass bei einzelnen (wenigen) Versuchspersonen die erhöhte Produktanzahl bereits zu einer deutlich längeren Entscheidungszeit geführt hat. Insgesamt sollte bei der Gestaltung von Designs also

darauf geachtet werden, die Anzahl an Auswahlobjekten so gering wie möglich zu halten, ohne dabei die Komplexität der Schnittstelle zu stark zu abstrahieren.

4.6 Das Paradox des aktiven Nutzers

Das Paradoxon des aktiven Nutzers beschreibt kurz zusammengefasst die Situation, dass aktive Nutzer eines Produktes oder Systems, zum Beispiel einer Software, in der Regel keine Instruktionen oder andere Hilfestellungen konsultieren, bevor sie mit ihrer tatsächlichen Arbeit beginnen. Auch Standardeinstellungen werden gemäss dem Paradoxon häufig nicht angepasst. Viel wahrscheinlicher ist es nach diesem Paradox, dass die Benutzer direkt beginnen, mit dem digitalen Produkt oder dem System zu arbeiten, um ihre Aufgaben zu erledigen, ohne davor die gesamte Arbeitsumgebung wahrzunehmen oder zu entdecken.

Die nachfolgend beschriebenen Experimente dienen der Untersuchung dieses Paradoxons. Auf Basis der Beobachtungen und Messungen soll empirisch bestimmt werden, inwiefern sich das theoretisch beschriebene Paradox in der Praxis bestätigen oder widerlegen lässt.

Es gilt zu beachten, dass es sich hier nicht um Experimente handelt, welche die Nutzer aktiv selbst durchführen, sondern um Beobachtungen, wie sich das Verhalten der Nutzer abhängig von unterschiedlichen Faktoren, wie dem Einblenden von Dialogen, Generieren von Feedback oder unterschiedlichen Informationsarchitekturen während der User Experience ändert oder beeinflussen lässt.

4.6.1 Testaufbau

Wirkung von Feedback – Teil 1

Die Probanden wurden während ihrer Interaktion mit der Webseite (User Experience) sowie während der Bearbeitung der Experimente beobachtet. Darüber hinaus wurden bestimmte Handlungen gezielt erfasst und dokumentiert. Dabei handelte es sich insbesondere um Aktionen wie das Anpassen von Standardeinstellungen oder das Aufrufen von Hilfestellungen und Instruktionen, wie sie bereits weiter oben beschrieben wurden.

Wie bereits erwähnt, wurden die Teilnehmenden in drei Gruppen (A, B und C) eingeteilt. Zu Beginn jeder User Experience mussten sich alle Studienteilnehmer mit ihrer E-Mail-Adresse im System anmelden. Nach erfolgreichem Anmeldeprozess wurden sie auf die Hauptseite der Webseite weitergeleitet. Ab diesem Zeitpunkt reagierte das System abhängig von der Gruppenzugehörigkeit unterschiedlich:

Für die Teilnehmenden der Gruppen A und B wurde automatisch ein Dialogfenster eingeblendet, das sie willkommen hiess und erste Informationen bereitstellte. Das Dialogfenster folgte dem Prinzip der Kontrolle: Es präsentiert nicht alle Inhalte auf einmal, sondern erlaubt es den Benutzern selbst zu entscheiden, wann sie welche Informationseinheit (Dialogseite) abrufen wollten. Über die Schaltflächen „Weiter“ und „Zurück“ können sie innerhalb der drei enthaltenen Informationseinheiten individuell navigieren (vgl. Anhang D, Paradox des aktiven Nutzers, Abbildung D.10 Dialogfenster).

Die erste Informationseinheit enthält allgemeine Hinweise zum Ablauf der Experimente, zum Aufbau und Inhalt der Webseite sowie zur Durchführung eines Versuchs. Die zweite Einheit informiert über Einstellungsmöglichkeiten, die eine schnellere Interaktion ermöglichen könnten. Diese Optionen betreffen insbesondere die oben erläuterte progressive Offenlegung der Experimentkategorien und der Versuchstests sowie die Funktion, automatisch den nächsten noch nicht abgeschlossenen Versuch innerhalb einer gewählten Experimentkategorie auszuwählen. Die dritte Informationseinheit weist auf die existierende Hilfeseite hin, die den Nutzern erklärt, wie sie innerhalb der Webseite navigieren, ein Experiment starten, welche Einstellungen zur Verfügung stehen und wie sich zusätzliche Tipps einblenden lassen. Abschliessend enthielt der Dialog den Hinweis, dass diese Informationen jederzeit während der Interaktion über einen blauen runden Button in der rechten unteren Ecke des Bildschirms erneut abgerufen werden können.

Die Gruppen unterschieden sich hinsichtlich der Art, wie der Dialog präsentiert wurde: Gruppe A konnte den Dialog jederzeit schliessen, auch unmittelbar nach dem Erscheinen. Die Teilnehmenden der Gruppe B mussten hingegen bis zur letzten Informationseinheit navigieren, bevor sie das Dialogfenster schliessen konnten. Bei der Gruppe C wurde der Dialog nicht automatisch angezeigt. Die Teilnehmer dieser Gruppe mussten den Informations-Button (ein blauer, runder Button in der rechten unteren Ecke der Webseite, vgl. Anhang D, Abbildung D.11) eigenständig erkennen und aktiv anklicken, um den Dialog zu öffnen.

Das Ziel dieses Versuchs ist es zu untersuchen, welchen Einfluss der beschriebene Dialog auf das Verhalten der Nutzer der jeweiligen Gruppen hat. Dabei sollten insbesondere folgende Fragen beantwortet werden:

- Wird Gruppe A den Dialog vollständig lesen oder ihn unmittelbar nach dem Erscheinen schliessen?
- Wie wird Gruppe B den Dialog durchgehen, wenn das vollständige Lesen erzwungen wird?

- Erkennt Gruppe C den Informationsbutton und nutzt diesen, um den Dialog freiwillig zu öffnen?
- Und schliesslich: Welche Teilnehmenden werden den Empfehlungen des Dialogs folgen, Einstellungen zur Beschleunigung der Interaktion anpassen oder die Hilfeseite konsultieren, anstatt bei Problemen direkt um Unterstützung zu bitten und die User Experience dadurch zu unterbrechen?

Wirkung von Feedback - Teil 2

Das zweite Experiment stellt eine Erweiterung des ersten dar und verfolgt das Ziel zu untersuchen, inwieweit das aktive Verhalten der Nutzer durch einen zusätzlichen Dialog, der den aktiven Interaktionsfluss gezielt unterbricht, beeinflusst wird. Das Experiment ist wie folgt aufgebaut: Sobald ein angemeldeter Nutzer nach Abschluss der zweiten Experimentkategorie zur Hauptseite der Webseite zurückkehrt und dabei die Einstellung der progressiven Offenlegung für die Experimentkategorien noch nicht aktiviert hat, so wird ein weiterer Dialog eingeblendet. Dieser weist darauf hin, dass durch die Aktivierung dieser Einstellung der Interaktionsprozess effizienter gestaltet und beschleunigt werden könnte. Um den Dialog zu schliessen, ist ein expliziter Mausklick auf einen Bestätigungsbutton erforderlich (vgl. Anhang D, Abbildung D.13). Wie in Kapitel 2, unter Abschnitt 2.3.3 erläutert wurde, handelt es sich bei dieser Art von Rückmeldung um ein sogenanntes modales Feedback, das den Interaktionsfluss der Nutzer gezielt unterbricht und ihre unmittelbare Aufmerksamkeit sowie eine Handlung erfordert. Das Ziel dieses Experiments besteht in der Analyse, ob diese Form des Feedbacks die Nutzer tatsächlich dazu motiviert, die angebotenen Einstellungen zu ändern oder ob der Dialog vielmehr ignoriert wird und unmittelbar geschlossen wird.

Leseverhalten der Nutzer

In einer weiteren Untersuchung geht es um die Analyse, wie sich das aktive Leseverhalten im Verlauf der User Experience der Nutzer in der Zeit verändert. Vor jedem Start eines Versuches wird dem Benutzer eine Versuchsbeschreibung visualisiert. Die Versuchsbeschreibung beschreibt im Detail, wie der Versuch funktioniert und was die Aufgaben der Benutzer sind. Gemäss dem Paradoxon des aktiven Nutzers ist davon auszugehen, dass die Nutzer die Versuchsbeschreibung nur bedingt, unvollständig oder unaufmerksam lesen. Je länger die Interaktion dauert, ist es möglich, dass sich dieses Verhalten sogar noch verstärkt. Das Experiment hat das Ziel, diese Annahme empirisch und qualitativ zu belegen oder zu widerlegen. Hierfür wird für jeden Versuch bestimmt, wie viel Zeit die Benutzer für das Lesen

der Versuchsbeschreibung investierten. Wie bereits beschrieben wurde, besitzt jeder aktiv laufende Versuch eine Versuchsanleitung (modeless Dialog), welche während der Durchführung der Versuche zu jeder Zeit konsultiert werden kann. Die Ergebnisse dieses Experimentes können Aufschluss darüber geben, ob solch vorab zur Verfügung gestellte Anleitungen oder Instruktionen den Nutzern eines Produktes helfen, ihre Ziele langfristig oder kurzfristig effizienter und mit weniger Aufwand zu erreichen, oder ob diese nur sporadisch oder gar nicht konsultiert werden aufgrund einer aktiven Benutzerführung, die zu jeder Zeit zugänglich ist, wie zum Beispiel die bereits erwähnte Versuchsanleitung, die zu jeder Zeit während eines Versuches auf- oder zugeklappt werden kann.

4.6.2 Usability- und User Experience-Metriken

Um das Nutzerverhalten während der User Experience zu messen, werden unterschiedliche Performance-Metriken verwendet und Aktionen der Probanden gezielt gemessen. Das Messen von Aktionen durch Nutzer kann als Issue-bezogene Metrik behandelt werden. Um das aktive Verhalten der Nutzer zu messen und anschließend zu analysieren, sind einerseits Zeitmetriken (Auftragszeiten), aber auch Self-reported-Metriken nützlich. Die Zeit, welche investiert wurde, den Dialog zu lesen, kann zum Beispiel Aufschluss darüber geben, wie fokussiert und aktiv die Informationen gelesen wurden. Das Messen von bestimmten Aktionen, wie zum Beispiel das Aufrufen eines Links oder das Klicken einer Schaltfläche, erlaubt es, das Verhalten der Nutzer weiter zu analysieren. Hierbei handelt es sich eher um Issue-bezogene Metriken. Nachbefragungen im Anschluss an die User Experience zum Verhalten der Nutzer, also Self-reported-Metriken, können schliesslich tiefere Einblicke in die Beweggründe der ausgeführten Handlungen der jeweiligen Nutzer geben und helfen, Usability-Issues zu identifizieren.

4.6.3 Ergebnisse

Wirkung von Feedback - Teil 1

Tabelle 4.10 zeigt die Anzahl der Teilnehmer der Gruppe A, die jeweils eine bestimmte Seite im ersten präsentierten Dialog (Willkommens-Dialog) erreichten. Zwei Teilnehmer schlossen den Dialog „blind“ ab, also unmittelbar nachdem er präsentiert wurde. Diese Teilnehmer lasen den Dialog also nicht. Die übrigen Teilnehmer aus Gruppe A erreichten dagegen alle die dritte und letzte Seite des Dialogs. Abbildung 4.10 zeigt die mittlere Lesezeit für den gesamten Dialog, der nach dem Authentifizierungsprozess visualisiert wurde, und die durchschnittliche Anzahl Se-

kunden, die Nutzer auf jeder einzelnen Seite verbrachten. Die Abbildung zeigt, dass die Benutzer der Gruppen A und B einige Sekunden investierten, den Dialog zu lesen. Gruppe B zum Beispiel las den Dialog im Schnitt beinahe eine Minute lang.

Seitenindex	Anzahl Teilnehmer
1	2
2	12
3	12

Tabelle 4.10: Gelesene Seiten im Willkommens-Dialog durch die Teilnehmer der Gruppe A

Tabelle 4.11 zeigt Extremwerte einzelner Teilnehmer. Wie der Tabelle zu entnehmen ist, investierten vier Studienteilnehmer zum Beispiel sogar mehr als 70 Sekunden, um den Dialog zu lesen. Ein Proband schloss den Dialog sogar erst nach 90 Sekunden. Die Teilnehmer der Gruppe C dagegen wiesen gemäss Abbildung 4.10 ein eher inaktives Verhalten auf. Kein Proband dieser Gruppe betätigte den runden Informationsbutton in der rechten unteren Ecke der Webseite, um den Dialog zu öffnen und dessen Informationen zu lesen.

Tabelle 4.12 veranschaulicht einige Aspekte des Nutzerverhaltens unmittelbar nach dem Schliessen des Dialogs sowie während des gesamten Interaktionsprozesses, also im Verlauf der gesamten User Experience. Obwohl die Teilnehmenden der Gruppen A und B nach Abbildung 4.10 ein eher aktives Leseverhalten zeigen, macht die Tabelle deutlich, dass das durch den Dialog vermittelte modale Feedback weder kurzfristig noch langfristig Einfluss auf das Verhalten der Nutzer hatte – unabhängig von ihrer Gruppenzugehörigkeit. Kein Teilnehmer folgte den im Dialog vorgeschlagenen Handlungen, wie etwa dem Aufrufen der Einstellungen zur Anpassung von Optionen. Nur ein Proband rief die Hilfeseite auf und ebenfalls nur ein Teilnehmer betätigte den runden Informationsbutton.

Wirkung von Feedback - Teil 2

Abbildung 4.11 zeigt, wie sich das Interaktionsverhalten einiger Nutzer nach dem Einblenden und Schliessen des zweiten Dialogs (Einblenden des Tipps) nach Abschluss der zweiten Experimentkategorie geändert hat. Wie die Grafik zeigt, hatte dieser Dialog eine Wirkung auf das aktive Verhalten mehrerer Nutzer, welche die progressive Offenlegung nach Abschluss der zweiten Kategorie nicht aktiviert hatten. Der Dialog veranlasste sieben Teilnehmer aus Gruppe A, die Einstellungen

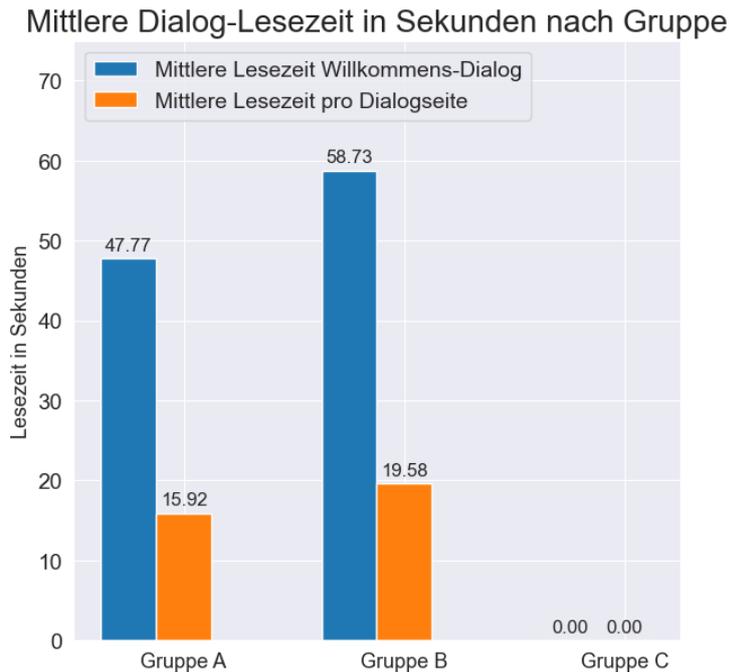


Abbildung 4.10: Mittlere Lesezeit des Willkommens-Dialog zu Beginn der User Experience jeder Gruppe

Teilnehmer	Anzahl Sekunden	Gruppe
1	90	A
2	80	B
3	76	B
4	73	B

Tabelle 4.11: Lesezeit des Dialoges einiger Teilnehmer in Sekunden

aufzurufen und die progressive Offenlegung zu aktivieren. Auch vier Teilnehmer der Gruppe B aktivierten diese Einstellung unmittelbar, nachdem sie den Dialog gelesen hatten. In Gruppe C schliesslich führte der Dialog bei drei Teilnehmern dazu, dass sie die Einstellungen aufriefen und die progressive Offenlegung aktivierten. Nur vier Teilnehmer von insgesamt 37 Teilnehmern aktivierten dagegen die automatische Auswahl des jeweils nächsten Versuches.

Leseverhalten der Nutzer

Als nächstes werden die Ergebnisse zur Analyse des Leseverhaltens der Nutzer präsentiert. Grafik 4.12 stellt links den Verlauf der mittleren Lesezeiten der Versuchsanleitungen in Sekunden für jede Experimentkategorie dar. Rechts dagegen

Aktion	Anzahl Teilnehmer
Unmittelbares Navigieren zu den Einstellungen	0
Betätigung des Informationsbutton	1
Hilfeseite aufrufen	1

Tabelle 4.12: Aspekte des Interaktionsverhaltens und unmittelbares Verhalten der Nutzer nach dem Schliessen des Willkommens-Dialog

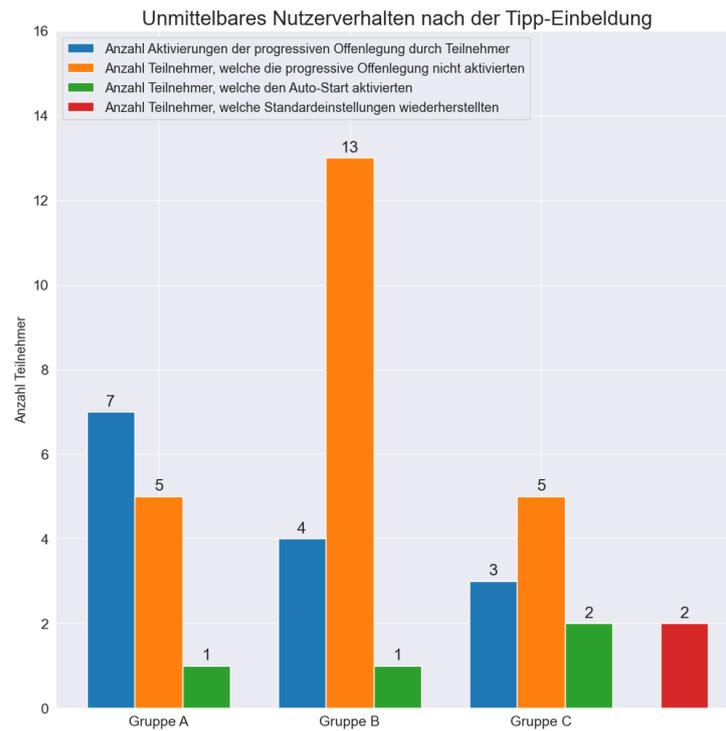
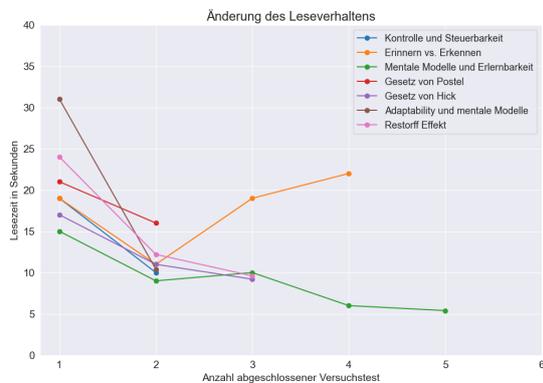
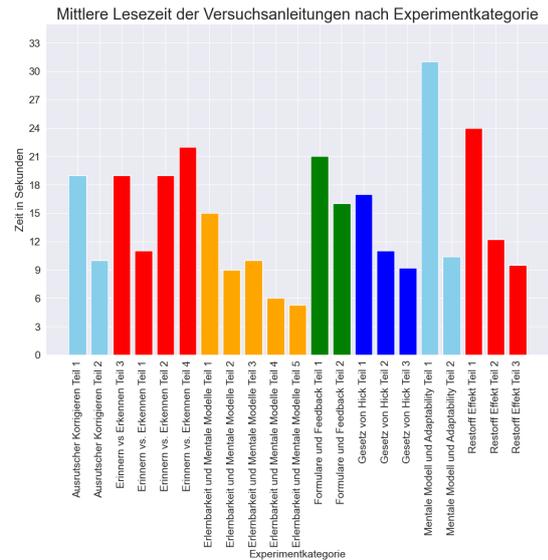


Abbildung 4.11: Unmittelbares Nutzerverhalten nach dem Schliessen des zweiten Dialoges nach Gruppe

wird für jeden einzelnen Versuchstest innerhalb einer Experimentkategorie die mittlere Lesezeit über alle Teilnehmer hinweg in Sekunden dargestellt. Dieselbe Farbe bedeutet dabei die Zugehörigkeit zu einer bestimmten Kategorie. Die beiden Grafiken zeigen einen deutlichen Trend im Leseverhalten der Nutzer: Je mehr Versuche innerhalb einer Kategorie abgeschlossen wurden, umso weniger Zeit wurde in das Lesen der Versuchsanleitungen investiert. Die einzige Ausnahme hier bildet die Versuchskategorie „Erinnern vs. Erkennen“. Wurde sodann eine neue Kategorie erreicht, so wurde die Versuchsanleitung des ersten Versuches aus dieser Kategorie wieder aufmerksamer während einer längeren Zeitspanne gelesen. Das aktive Verhalten der Nutzer pendelt also je nach Interaktionsfortschritt hin und her.



(a) Entwicklungen der mittleren Lesezeiten der Versuchsanleitungen innerhalb jeder Kategorie



(b) Mittlere Lesezeiten der Versuchsanleitungen jeder Experimentkategorie

Abbildung 4.12: Ergebnisse Leseverhalten: Änderungen der investierten Zeit um die Versuchsanleitungen zu lesen

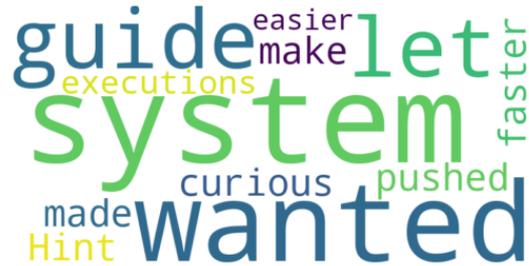
4.6.4 Interpretation der Ergebnisse

Nachbefragung

Um mögliche Beweggründe für das beobachtete Nutzerverhalten zu finden, wurden die Studienteilnehmer nach der Durchführung des Usability-Tests zu ihrem Verhalten unmittelbar nach dem Schliessen des Willkommensdialogs befragt. Mit Rückblick auf die beobachteten Verhaltensweisen, welche die Ergebnisse oben darstellen, wurden die Teilnehmer gefragt, warum sie nicht den Vorschlägen des Dialogs folgten und die entsprechenden Handlungen ausführten. Die genannten Gründe sind vielfältig und in Grafik 4.13 durch eine Wortwolke dargestellt. Wie aus der Grafik 4.13a abgelesen werden kann, liegen zentrale Gründe an der Menge an Informationen, welche auf einmal gelesen und verarbeitet werden mussten und offenbar nicht im Kurzzeitgedächtnis gespeichert werden konnten sowie am mangelnden Lesefokus der Nutzer. Diese Gründe sind nachvollziehbar, denn tatsächlich kommunizierte der Dialog ein umfangreiches Feedback, welches offenbar das Gesetz von Miller nicht genug berücksichtigt, da das Kurzzeitgedächtnis von den Probanden aufgrund der Menge an Informationen zu stark beansprucht wird. Weiter kann hier auch das Prinzip der Kontrolle und Steuerbarkeit in Betracht gezogen werden. Wird nämlich die progressive Offenlegung aktiviert, wird Kontrolle an das System abgegeben, was dazu führt, dass die individuelle Gestaltung der User Experience stark eingeschränkt wird. Einige Teilnehmer waren nicht bereit, diese Kontrolle abzugeben.



(a) Verteilung der mittleren Benutzer-Lostness-Werte



(b) Verteilung der maximalen Lostness-Wertebereiche der Teilnehmer

Abbildung 4.13: Äusserungen der Nutzer zu ihrem unmittelbaren Verhalten nach dem Schliessen des ersten und zweiten Dialogs

Abbildung 4.13b präsentiert die Ergebnisse der Nachbefragungen zum unmittelbaren Verhalten der Teilnehmenden, nachdem sie den zweiten Hinweisdialog geschlossen haben. Diese Ergebnisse präsentieren nur Antworten von Teilnehmern, welche unmittelbar nach dem Bestätigen des Dialogs die Einstellungen aufrufen und änderten. Wie der Grafik entnommen werden kann, sind auch hier die Antworten der Teilnehmer vielseitig. Einige der Probanden äusserten, dass sie sich vom System gedrängt fühlten, die Einstellungen aufzurufen, andere dagegen waren bloss neugierig, was die progressive Offenlegung bewirkt. Widerum andere wollten den Aufforderungen des Systems folgen und wie vom System mitgeteilt, den Interaktionsprozess schneller oder einfacher abschliessen. Mit Ausnahme von zwei Teilnehmenden haben alle Probanden den Hinweis-Dialog nicht unmittelbar nach dessen Erscheinen geschlossen. Ein möglicher Grund dafür könnte die Informationsmenge im Dialogfenster sein. Der angezeigte Text war kurz und prägnant und wurde vermutlich nicht als störend oder ablenkend wahrgenommen.

Leseverhalten

Das in Abbildung 4.12 erkennbare Verhaltensmuster beim Lesen der Versuchsanleitungen lässt sich möglicherweise wie folgt erklären: Vor dem erstmaligen Bearbeiten eines Versuchs innerhalb einer neuen Kategorie bestand bei den Nutzern vermutlich ein verstärktes Interesse daran, die Zielsetzung und die wesentlichen Aufgaben des Versuchs zu verstehen. Im weiteren Verlauf der Bearbeitung der übrigen Versuche innerhalb derselben Kategorie lernten und erinnerten sich die Teilnehmer zunehmend, dass die grundlegende Aufgabe innerhalb einer Kategorie ähnlich bleibt und sich lediglich die Bedienoberfläche verändert. Daher investierten sie mit jeder weiteren Aufgabe vermutlich weniger Zeit in das Lesen der Versuchsanleitung. Ein weiterer möglicher Einflussfaktor auf dieses Verhalten könnte ein Erinnerungs- und Lernprozess sein: Während der Experimente wurde beobach-

tet, dass bereits beim ersten Versuch einer Versuchsreihe die Nutzer die dauerhaft sichtbare Schritt-für-Schritt-Anleitung eines Versuchs erkannten, die während der gesamten Durchführung präsent war und sich dynamisch anpasste. Noch bevor sie die Versuchsanleitung des jeweils nächsten Versuchs aufrufen, erinnerten sich die Teilnehmer wahrscheinlich an diese unterstützende Anleitung. In dem Wissen, dass sie erneut schrittweise durch den Versuch geführt werden würden, neigten viele dazu, die detaillierte Versuchsanleitung nur noch oberflächlich zu überfliegen.

4.7 Gesetz von Tesler

Um die Auswirkungen von Teslers Gesetz auf die Usability und User Experience der Nutzer zu untersuchen, wird in den folgenden Experimenten gezielt das zuvor erläuterte Prinzip der progressiven Offenlegung eingesetzt (vgl. Kapitel 2, Abschnitt 2.3.1, Progressive Offenlegung). Die konkrete Umsetzung dieser Funktion als optionale Einstellung innerhalb der Webseite wurde bereits in diesem Kapitel unter Abschnitt 4.2.2, Einstellungen, beschrieben.

Kurz zusammengefasst besagt das Gesetz von Tesler, dass jeder Interaktionsprozess einen unvermeidbaren, minimalen Komplexitätsgrad aufweist, der nicht weiter reduziert oder abstrahiert werden kann. Ziel des folgenden Experiments ist es zu analysieren, in welchem Ausmass eine gezielte Benutzerführung durch das schrittweise Einblenden von Aktions- und Auswahllementen, also die progressive Offenlegung, die Effizienz der Ausführung sowie der Entscheidungsfindung während des Interaktionsprozesses beeinflusst.

4.7.1 Testaufbau

Reduzierung des Komplexitätsgrades

Die Aufgabe der Probanden während des Usability-Tests und der User Experience bestand darin, alle auf der Webseite verfügbaren Versuchsexperimente durchzuführen. Standardmässig können die Teilnehmer dabei frei entscheiden, in welcher Reihenfolge sie die Experimente absolvieren wollen. Das bedeutet, dass die Probanden vor jedem neuen Versuch selbst wählen, aus welcher Kategorie sie einen Versuch durchführen. Zu beachten ist hierbei nun, dass alle Versuchsexperimente gleichwertig sind, wodurch der Entscheidungsprozess bezüglich der Auswahl einer Kategorie und eines Versuchs theoretisch keine Gewichtung durch den Nutzer erfordert. Die Komplexität dieses Entscheidungsprozesses lässt sich im Vergleich zur beschriebenen Standardsituation gezielt reduzieren und abstrahieren, indem dem Nutzer bei jedem Auswahlprozess jeweils nur eine noch nicht abgeschlossene Experimentkate-

gorie beziehungsweise ein noch offener Versuch präsentiert wird. Diese Reduktion der Wahlmöglichkeiten wird durch die Aktivierung der progressiven Offenlegung erreicht. Gleichzeitig wird dadurch eine indirekte Form der Benutzerführung in den Ablauf integriert. Die Wirkung der progressiven Offenlegung auf die User Experience wird untersucht, indem der beschriebene Entscheidungsprozess jener Probanden, die diese Funktion während des gesamten Interaktionsprozesses deaktiviert liessen, mit dem Entscheidungsverhalten jener Teilnehmer verglichen wird, die die progressive Offenlegung im Verlauf der User Experience aktivierten.

4.7.2 Usability und User Experience-Metriken

Performance-Metriken

Zur Evaluierung der Auswirkungen der progressiven Offenlegung auf die Performance der Nutzer werden verschiedene Performance-Metriken herangezogen. Performance-Metriken erlauben es, die Effizienz und den Aufwand der Handlungsprozesse der Nutzer zu erfassen, also die Leistungsfähigkeit und Effizienz der Nutzer während der Entscheidungsfindung und Navigation zu messen und zu beurteilen.

Ein zentraler Wert stellt dabei die benötigte Zeit dar, die Teilnehmende aufwenden, um eine Entscheidung darüber zu treffen, welche Kategorie oder welcher Versuch innerhalb einer Kategorie als nächstes ausgewählt werden soll. Diese Zeitangabe liefert Aufschluss über die Entscheidungseffizienz der Nutzenden. Ergänzend dazu erscheint es sinnvoll, gegebenenfalls das Entscheidungsverhalten einzelner Personen näher zu betrachten, insbesondere wenn deren Vorgehen deutlich vom Durchschnitt abweicht.

Ziel der Analyse mit unter Verwendung dieser Metriken ist es zu untersuchen, ob die progressive Offenlegung zu einer beschleunigten und insgesamt effizienteren Interaktion beiträgt und inwiefern eine aktive Benutzerführung die kognitive Belastung der Nutzenden bei Entscheidungsprozessen reduzieren kann.

4.7.3 Ergebnisse

Abbildung 4.14 zeigt die mittlere Auswahlzeit der Benutzer, welche die progressive Offenlegung während der User Experience aktivierten und jener Teilnehmer, welche dies nicht taten. Die Grafik zeigt, dass die Nutzer, welche die progressive Offenlegung im Verlauf der Interaktion aktivierten, im Schnitt eine geringere Entscheidungszeit aufweisen, als die Teilnehmergruppe, welche diese Einstellung nie aktivierte.

Grafik 4.15 verdeutlicht beispielhaft an vier Teilnehmern, was passieren kann,



Abbildung 4.14: Mittlere Auswahlzeit der Nutzer für Experimentkategorien

wenn die Komplexität eines Entscheidungsprozesses nicht reduziert, resp. nicht an das System übergeben wird. Wie die Abbildung zeigt, führte die fehlende progressive Offenlegung bei zwei Teilnehmern zu einer sehr hohen Entscheidungszeit. Beide Teilnehmer benötigten jeweils mehr als eine Minute, um sich für die erste Kategorie zu entscheiden. Dieses Verhalten kann im Kontext der im Rahmen der Arbeit durchgeführten User Experience bereits als Usability-Issue identifiziert werden. Dagegen wiesen die zwei anderen Teilnehmer, welche die progressive Offenlegung aktivierten, ein viel effizienteres Interaktionsverhalten auf.

Als nächstes soll ein Blick auf die Entscheidungseffizienz der Nutzer im Verlauf der User Experience geworfen werden. Dazu eignet sich die Betrachtung der Abbildung 4.16 und Abbildung 4.17. Abbildung 4.16a zeigt die Veränderung der mittleren Entscheidungszeit in Abhängigkeit von der Anzahl bereits gewählter Kategorien – getrennt nach Nutzern, die die Progressive Offenlegung (PD) aktiviert haben, und solchen, die dies nicht taten. Die rechte Abbildung (Abbildung 4.16b) veranschaulicht hingegen die Veränderung der mittleren Entscheidungszeit für Versuchsexperimente innerhalb der jeweiligen Kategorien, ebenfalls differenziert nach Aktivierung oder Deaktivierung der PD. Beide Darstellungen machen deutlich, dass die Nutzer, welche die PD im Verlauf der User Experience aktiviert haben, über nahezu den gesamten Prozess hinweg eine effizientere Entscheidungsfindung zeigten als diejenigen Teilnehmenden, die die PD nicht nutzten.

Auswahlzeit der ersten Kategorie für 4 Teilnehmer (T) mit/ohne progressive Offenlegung (PD)

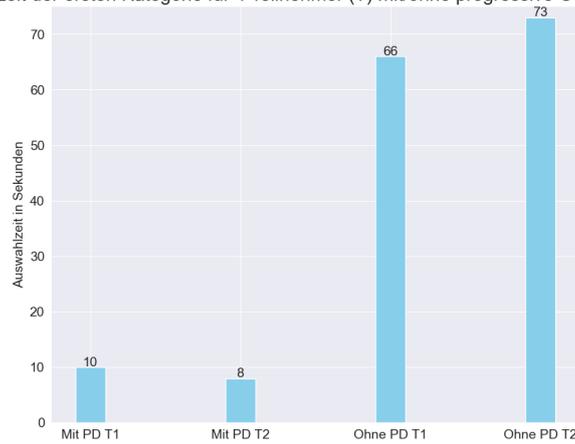
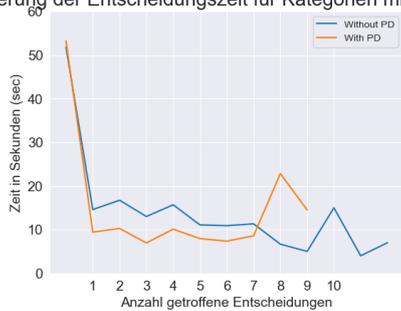
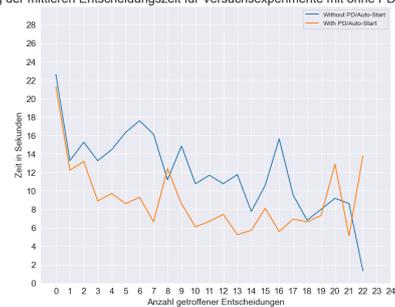


Abbildung 4.15: Auswahlzeiten der ersten Kategorie für vier Teilnehmer mit und ohne progressiver Offenlegung (PD)

Änderung der Entscheidungszeit für Kategorien mit/ohne PD



Änderung der mittleren Entscheidungszeit für Versuchsexperimente mit ohne PD oder Auto-Start

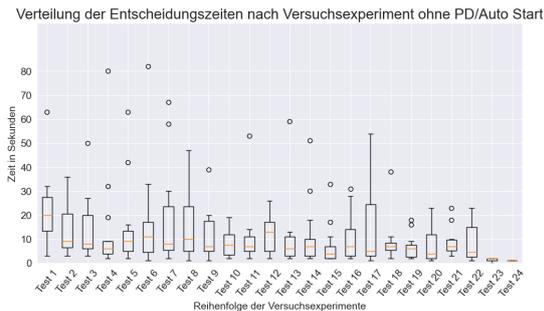


(a) Änderung der mittleren Entscheidungszeit für Kategorien mit und ohne Progressive Offenlegung oder Auto-Start

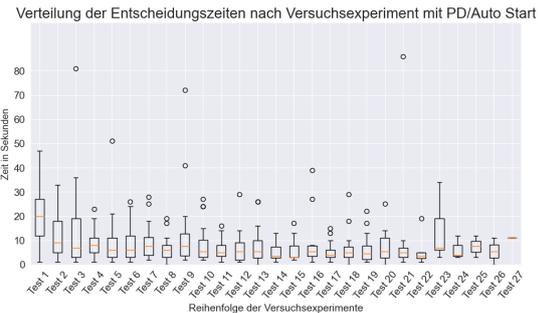
(b) Änderung der mittleren Entscheidungszeit für Versuchstests mit und ohne Progressive Offenlegung oder Auto-Start

Abbildung 4.16: Änderung der mittleren Entscheidungszeit für Kategorien und Versuchstests mit und ohne progressive Offenlegung (PD) oder Auto-Start

Auch Abbildung 4.17 stützt diese Aussage. Die Balken in Abbildung 4.17a, die teils weit nach oben reichen, während sich die braune Linie (Median) im unteren Bereich der Box befindet, deuten darauf hin, dass die höhere Komplexität des Entscheidungsprozesses – bedingt durch das Fehlen der PD – bei einigen Teilnehmenden zu einer geringeren Effizienz bei der Auswahl der Versuchsexperimente geführt hat. Im Gegensatz dazu zeigen die Balken in Abbildung 4.17b eine deutlich symmetrischere Verteilung der Entscheidungszeiten bei den Probanden, die die PD im Verlauf der User Experience aktiviert haben. Dies spricht dafür, dass es in dieser Gruppe weniger Teilnehmende mit stark überdurchschnittlichen Entscheidungszeiten gibt.



(a) Verteilung der Entscheidungszeiten nach Kategorie ohne Progressive Offenlegung oder Auto-Start



(b) Verteilung der Entscheidungszeiten nach Versuchstest mit Progressive Offenlegung oder Auto-Start

Abbildung 4.17: Änderung der mittleren Entscheidungszeit für Kategorien und Versuchstests mit und ohne progressive Offenlegung oder Auto-Start

4.7.4 Interpretation der Ergebnisse

Bei der Interpretation der Ergebnisse im Zusammenhang mit Teslers Gesetz sollen keine übermäßigen Spekulationen betrieben werden. Vielmehr liegt es nahe, dass die Resultate eine direkte Folge dieses Gesetzes darstellen: Je geringer die Komplexität eines Prozesses – unter Berücksichtigung eines notwendigen minimalen Komplexitätsgrades – ausfällt, desto effizienter können Nutzer agieren und desto weniger mentale Ressourcen sind für die Bearbeitung des Prozesses erforderlich. Die Progressive Offenlegung reduzierte die Komplexität des Entscheidungsprozesses spürbar, während die automatische Auswahl des jeweils nächsten Versuchstests die Komplexität dieses Prozesses nahezu auf ein Minimum reduzierte. Die Ergebnisse (vgl. Abbildung 4.15) zeigen, dass bei einzelnen Individuen eine fehlende Entscheidungsübernahme durch das System bei nicht zwingend notwendigen Entscheidungen zu einem Usability-Issue führen kann, zum Beispiel eine sehr geringe Entscheidungseffizienz.

4.8 Kontrollierbarkeit und Steuerbarkeit

Um das Prinzip der Kontrollierbarkeit und Steuerbarkeit zu untersuchen, wurden zwei Versuchsvarianten entwickelt. In beiden Untersuchungen geht es konkret um die Analyse, wie Nutzer reagieren, wenn diese in die Situation gelangen, welche durch sie gegebenenfalls als Ausrutscher oder Fehler identifiziert wird. Gegenstand der Untersuchung ist, wie Nutzer Möglichkeiten identifizieren oder wahrnehmen, Ausrutscher zu beheben oder wie sie Handlungssequenzen steuern, um das System wieder in einen korrekten Zustand zu überführen.

4.8.1 Testaufbau

Im Rahmen der Untersuchung des Prinzips der Kontrolle und Steuerbarkeit wurden ähnliche, sich nur gering unterscheidende grafische Benutzeroberflächen einer E-Mail-Anwendung entwickelt (vgl. Anhang D, Kontrollierbarkeit und Steuerbarkeit, Abbildung D.16). Beide Versionen ermöglichen es den Nutzenden, E-Mails zu verwalten. Nach dem Start der jeweiligen Testvariante wird standardmässig der Posteingang der Anwendung angezeigt. Beide Oberflächen verfügen über eine Menüschnittstelle, über die gesendete, gelöschte, Spam-E-Mails oder archivierte E-Mails mittels entsprechender Links aufgerufen werden können. Diese Menükomponente ist als Dropdown-Menü gestaltet, trägt den Titel „Ordner“ und verbirgt die zugehörigen Links im Ausgangszustand. Ein Mausklick auf das Menü zeigt den Inhalt an oder blendet ihn wieder aus (vgl. Anhang D, Abbildung D.16 und D.17).

Zu Beginn jeder Testvariante informiert die Versuchsanleitung oben links in der Ecke des Bildschirms die Teilnehmenden darüber, welche E-Mail zu löschen ist. Die E-Mail, welche gelöscht werden soll, wird dabei zufällig durch das System bestimmt. Das Löschen erfolgt über eine Schaltfläche am rechten Rand jeder E-Mail. Nach dem Löschvorgang erscheint für acht Sekunden eine „Rückgängig“-Schaltfläche (Undo-Funktion), über die der Löschvorgang widerrufen werden kann. Acht Sekunden entsprechen zum Beispiel in etwa der gleichen Zeit, wie lange die E-Mail-Anwendung von Outlook³ die Undo-Funktion darstellt.

In der ersten Testvariante wird die Undo-Funktion in der Mitte am unteren Bildschirmrand dargestellt (vgl. Anhang D, Abbildung D.18). In der zweiten Variante befindet sie sich in der oberen linken Bildschirmecke oberhalb der Versuchsanleitung. Wird das Löschen nicht widerrufen oder das Menü unmittelbar nach dem Löschen betätigt, so wird die E-Mail, welche gelöscht wurde, am Anfang der Liste aller bisher gelöschten E-Mails hinzugefügt, wobei der Betreff, Absender und der gezeigte Textinhalt in der zweiten Versuchsvariante farblich hervorgehoben werden. Diese farbliche Visualisierung erfolgt in der ersten Variante nicht. Somit wird in der zweiten Variante gezielt der Restorff-Effekt eingesetzt. Wird die Rückgängig-Funktion genutzt, wird die gelöschte E-Mail an exakt derselben Position im Posteingang wieder eingefügt, aus der sie zuvor entfernt wurde. Auch hier unterscheiden sich die beiden Varianten: In der zweiten Variante wird die wiederhergestellte E-Mail zusätzlich farblich hervorgehoben, in der ersten Variante fehlt diese visuelle Markierung (vgl. Anhang D, Kontrollierbarkeit und Steuerbarkeit, Abbildung D.19).

Die anschliessenden Aufgaben variieren zwischen den beiden Testvarianten. In der ersten Variante werden die Nutzenden nach dem Löschen der E-Mail unmit-

³Microsoft 365, Outlook, abgerufen am 21.06.2025; <https://www.microsoft.com/de-ch/microsoft-365/outlook/email-and-calendar-software-microsoft-outlook?market=ch>

telbar aufgefordert, das Versanddatum der gelöschten Nachricht in ein Eingabefeld einzutragen. In der zweiten Variante sollen sie hingegen den Betreff der gelöschten E-Mail in das entsprechende Feld eingeben. Durch das unmittelbare Verlangen von Informationen von einem gelöschten Objekt wird absichtlich versucht, den Nutzern das Gefühl zu verleihen, dass ihnen ein Ausrutscher passierte. Die Idee ist die, dass sie die getätigte Aktion als Fehler interpretieren.

Ziel des Experiments ist es sodann, das Verhalten der Nutzenden nach dem Löschen der E-Mail zu analysieren. Untersucht wird, auf welche Weise sie versuchen, die geforderten Informationen – also Versanddatum oder Betreff der gelöschten E-Mail – zu beschaffen. Hierbei stehen ihnen zwei mögliche Vorgehensweisen zur Verfügung: Entweder sie nutzen die Undo-Funktion, um die E-Mail wiederherzustellen, oder sie greifen über den Ordner über den Link „gelöschte Elemente“ auf die benötigten Informationen zu, um die Aufgabe erfolgreich zu lösen.

4.8.2 Usability und User Experience-Metriken

Zur Erfassung des Nutzerverhaltens werden einerseits Performance-Metriken und andererseits subjektive Selbstberichte (Self-reported-Metriken) eingesetzt.

Issue-bezogene-Metriken

Als Issue-bezogene Metriken können die Häufigkeiten der jeweiligen Vorgehensweisen gezählt werden. Damit kann verglichen werden, wie viele Nutzer die Undo-Funktion nutzten und wie viele Probanden den alternativen Weg über die gelöschten E-Mails einschlugen. Zudem kann auf diese Weise evaluiert werden, welche Vorgehensweise aus Effizienz­sicht möglicherweise zu einem Usability-Issue führen kann.

Performance-Metriken

Performance-Metriken dienen dazu, die Effizienz der unterschiedlichen Vorgehensweisen – Nutzung der Undo-Funktion oder Zugriff auf den Link „Gelöschte E-Mails“ im Ordner – zu vergleichen. Entscheidet sich ein Nutzer dafür, die gelöschten E-Mails aufzurufen, wird die Zeit gemessen, die vergeht, bis der entsprechende Link im Drop-Down-Menü erkannt wird. Diese Messung kann als Indikator für Erkennbarkeit und Erlernbarkeit der Benutzeroberfläche interpretiert. Als weitere Performance-Metrik wird auch die Erfolgsquote betrachtet. Dabei wird gemessen, ob die Teilnehmenden die Informationen der korrekten, tatsächlich gelöschten E-Mail eingegeben haben. Für diejenigen Nutzer, welche die Undo-Funktion einsetzen, werden auch die Reaktionszeiten analysiert. Das ist die Zeit in Sekunden, welche die

Nutzer benötigten, um die Undo-Funktion zu betätigen, ab dem Moment, wo sie visualisiert wurde.

Self-reported-Metriken

Um die beiden beschriebenen Versuchsvarianten erfolgreich abzuschliessen, müssen die Nutzer die Funktionsweise erkennen und erlernen. Self-reported-Metriken erlauben es dabei, die wahrgenommene kognitive Belastung und die geschätzte Erlernbarkeit der beiden Versuchsvarianten zu erfassen. Dabei handelt es sich um subjektive Einschätzungen der Teilnehmenden hinsichtlich des empfundenen mentalen Aufwands, um die jeweilige Aufgabe zu bewältigen. Ziel ist es, Unterschiede in der kognitiven Beanspruchung zwischen den beiden Interaktionsvarianten zu identifizieren und zu vergleichen.

Kombinierte und vergleichende Metriken

Als letzte Art von Metriken sollen auch kombinierte und vergleichende Metriken herangezogen werden. Es bietet sich hier an, Self-reported-Metriken, also subjektive Bewertungen der kognitiven Belastung mit der Auftragszeit zu kombinieren, um zu untersuchen, ob eine längere Bearbeitungszeit tendenziell zu einer höheren kognitiven Belastung führt.

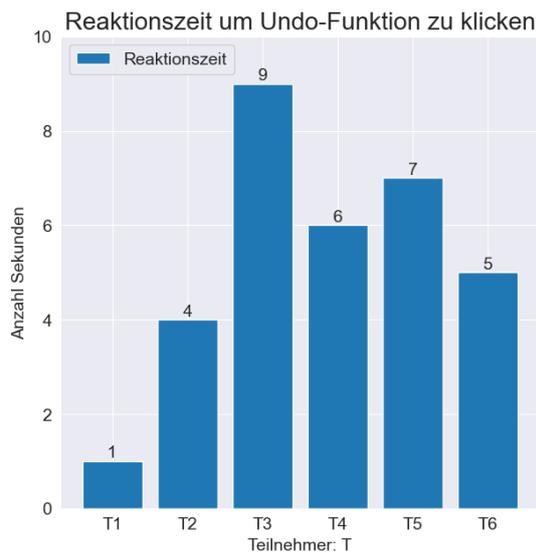
4.8.3 Ergebnisse

Nachfolgend werden nun die Untersuchungsergebnisse zu den beiden Versuchsvarianten präsentiert.

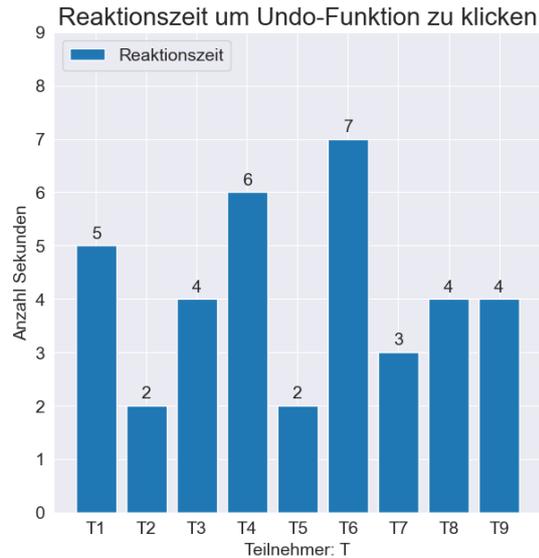
Wie aus der Tabelle 4.13 gelesen werden kann, hat eine Mehrheit der Nutzer die Undo-Funktion nicht eingesetzt, resp. nicht erkannt, um die E-Mail wiederherzustellen und auf diese Art und Weise an die verlangten Informationen zu gelangen. Die meisten Probanden, 28 im ersten und 25 im zweiten Versuch, haben den Weg über die gelöschten E-Mails eingeschlagen. Das bedeutet, sie haben nach dem Löschen der E-Mail und dem Lesen der Instruktion in der Versuchsanleitung den Menü-Ordner gesucht und danach den Link „gelöschte Elemente“ aufgerufen, um die gelöschten E-Mails abzurufen, um dann die gelöschte E-Mail zu lokalisieren und die Informationen, welche benötigt wurden, einzutippen.

Versuchsnummer	Anzahl Teilnehmer mit Undo-Funktion	Anzahl Teilnehmer ohne Undo-Funktion
1	6	28
2	9	25

Tabelle 4.13: Anzahl Teilnehmer, welche die Undo-Funktion einsetzen/nicht einsetzen, um die gelöschte E-Mail wiederherzustellen



(a) Reaktionszeit, um die Undo-Funktion zu betätigen im ersten Versuch



(b) Reaktionszeit, um die Undo-Funktion zu betätigen im zweiten Versuch

Abbildung 4.18: Reaktionszeiten der Nutzer, um die Undo-Funktion im ersten und zweiten Versuch zu betätigen

Reaktionszeiten

Abbildung 4.18 zeigt die Reaktionszeiten in Sekunden der einzelnen Teilnehmer, die die Undo-Funktion verwendet haben. Die Reaktionszeit beschreibt dabei den Zeitraum zwischen dem Einblenden der Undo-Funktion und deren Betätigung. Da nur wenige Teilnehmende die Funktion genutzt haben, werden in der Abbildung die individuellen Reaktionszeiten der einzelnen Personen dargestellt. Wie ersichtlich ist, variieren die Reaktionszeiten stark – sie reichen von einer bis zu neun Sekunden. Mit Ausnahme des ersten Teilnehmers in Abbildung 4.18a sowie der Teilnehmenden zwei und fünf in Abbildung 4.18b zeigen die übrigen Probanden hohe Reaktionszeiten. Dies deutet darauf hin, dass die Undo-Funktion in beiden Versuchsvarianten häufig nicht unmittelbar nach dem Einblenden erkannt wurde.

Auftragszeit

Als nächstes werden die Bearbeitungszeiten der Nutzer für die beiden Versuchsvarianten betrachtet. Es erfolgt ein Vergleich der Auftragszeiten derjenigen Nutzer, welche die Undo-Funktion nutzen, mit denjenigen der Teilnehmenden, welche die Undo-Funktion nicht verwendeten. Diese Ergebnisse werden von Abbildung 4.19 präsentiert. Diese stellt die mittleren Bearbeitungszeiten aller Teilnehmenden je Versuch dar. Dabei wird auch unterschieden, ob die Undo-Funktion eingesetzt wurde oder nicht. Wie in der Tabelle zu erkennen ist, benötigten die Probanden, welche

die Undo-Funktion nicht nutzen, im Schnitt mehr Zeit als die Teilnehmer, welche die Undo-Funktion einsetzen. Es muss dabei aber berücksichtigt werden, dass es viel mehr Nutzer gab, welche die Undo-Funktion nicht einsetzen, als solche, welche sie verwendeten.

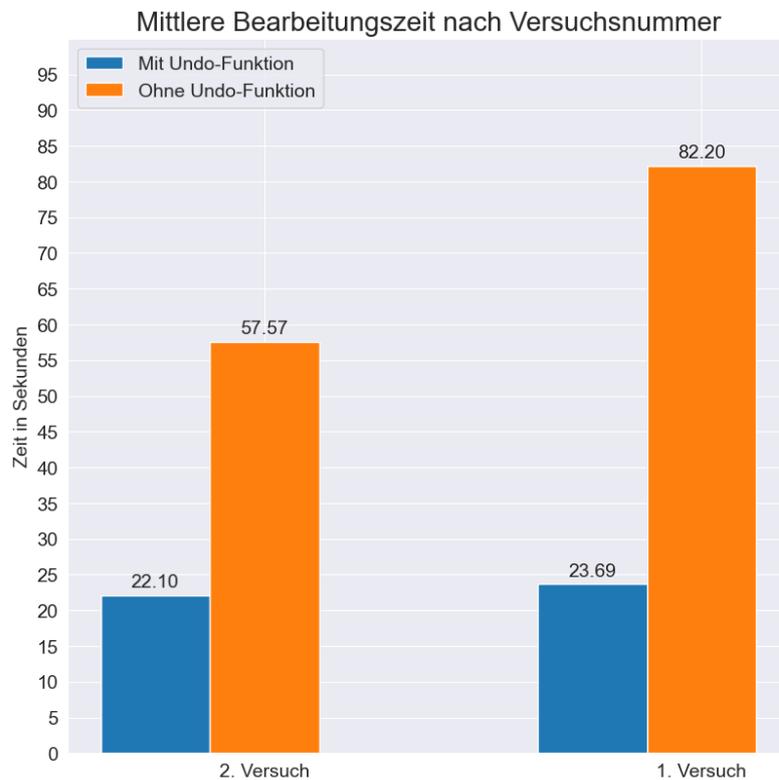


Abbildung 4.19: Mittlere Bearbeitungszeiten nach Versuchsnummer und Einsatz der Undo-Funktion

Einprägsamkeit und Erlernbarkeit

Da für die Untersuchung des Prinzips der Kontrolle und Steuerbarkeit zwei Versuche entwickelt wurden, in denen ein Einprägungs- und Lernprozess zu einer höheren Effizienz im zweiten Versuch führen kann, werden hier die Ergebnisse der einzelnen Nutzer aus beiden Versuchen verglichen. Dies erlaubt es, Aussagen zur Lernbarkeit und Einprägsamkeit zu treffen. Zunächst werden die Auftragszeiten der Nutzer aus beiden Versuchen verglichen. Diese Ergebnisse liefert Tabelle 4.14. Die Tabelle zeigt die beiden Bearbeitungszeiten für die Teilnehmer, welche die Undo-Funktion nicht einsetzen. Wie zu erkennen ist, konnten einige Teilnehmer ihre Bearbeitungszeit im zweiten Versuch im Vergleich zum ersten Versuch reduzieren, wurden also effizienter. Insgesamt 14 Teilnehmer, also 50%, konnten ihre Interaktionsgeschwindigkeit erhöhen. Auf der anderen Seite jedoch gab es solche Teilnehmer, welche ihre Interaktionsgeschwindigkeit nicht erhöhen konnten, die also mehr Zeit brauchten, den

zweiten Versuch abzuschliessen.

Teilnehmer	Auftragszeit (Seckunden) 1. Versuch	Auftragszeit (Sekunden) 2. Versuch
1	42.52	43.99
2	45.04	51.33
3	39.23	79.99
4	29.98	37.69
5	22.87	23.23
6	117.70	56.53
7	100.39	59.75
8	155.72	104.61
9	129.58	43.20
10	40.80	111.03
11	79.29	138.79
12	65.99	92.67
13	169.83	79.78
14	87.69	77.96
15	52.12	43.64
16	20.50	24.96
17	63.92	70.82
18	61.31	36.52
19	62.55	44.52
20	77.11	34.27
21	58.79	43.70
22	246.79	27.55
23	33.76	42.57
24	72.46	38.90
25	77.96	31.27
26	159.78	–
27	105.37	–
28	82.63	–

Tabelle 4.14: Bearbeitungszeit nach Versuchsnummer und Teilnehmer, welche die Undo-Funktion nicht einsetzen

Tabell 4.15 präsentiert die Auftragszeiten der Nutzerinnen und Nutzer in Sekunden, welche in beiden Versuchsvarianten die Undo-Funktion verwendeten. Von diesen insgesamt sechs Personen konnten vier davon ihre Interaktionsgeschwindigkeit im zweiten Versuch im Vergleich zum ersten beschleunigen.

Teilnehmer	Auftragszeit (Sekunden) 1. Versuch	Auftragszeit (Sekunden) 2. Versuch
1	24.87	14.69
2	18.67	20.87
3	21.41	17.13
4	21.71	35.70
5	26.53	18.26
6	28.92	21.13

Tabelle 4.15: Bearbeitungszeiten der Teilnehmer, welche die Undo-Funktion einsetzen

Tabelle 4.16 zeigt schliesslich noch die mittlere Anzahl Sekunden an, die die Teilnehmer in beiden Versuchen benötigten, um den Ordner mit dem Link „gelöschte Elemente“ zu finden und anzuklicken. Wie deutlich zu erkennen ist, konnten die

Teilnehmer diese Anzahl Sekunden im zweiten Versuch im Vergleich zum ersten reduzieren. Sie erkannten also den Ordner und den zielführenden Link im zweiten Versuch im Schnitt deutlich schneller als im ersten Versuch.

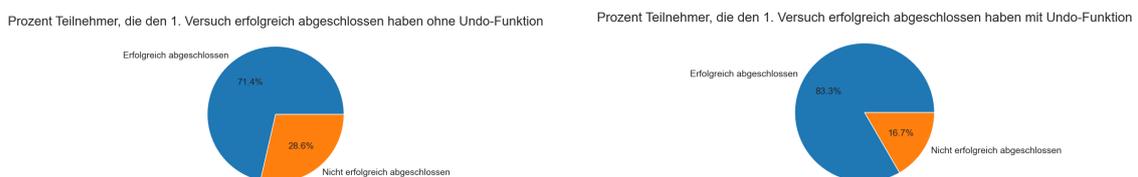
Versuchsnummer	Mittlere Anzahl Sekunden um den Link „gelöschte Elemente“ zu finden
1	51.2
2	36.86

Tabelle 4.16: Mittlere Zeit um den Ordner mit dem Eintrag „gelöschte Elemente“ zu finden und zu klicken

Auftragserfolg

Da in beiden Versuchsvarianten das Erreichen des Endzustandes nicht automatisch bedeutet, dass die Aufgabe erfolgreich abgeschlossen wurde, werden nachfolgend die Auftragserfolgsquoten präsentiert. Beide Versuche gelten dann als erfolgreich abgeschlossen, wenn die Informationen der tatsächlich gelöschten E-Mails korrekt ins Eingabefeld eingegeben worden sind.

Abbildung 4.20 zeigt den Anteil an Probanden, welche den ersten Versuch erfolgreich resp. nicht erfolgreich abgeschlossen haben, unterteilt in diejenigen Nutzer, welche die Undo-Funktion einsetzen (rechts) und diejenigen, welche sie nicht einsetzen (links).



(a) Verteilung der Teilnehmer, welche den 1. Versuch erfolgreich/nicht erfolgreich und ohne Undo-Funktion bearbeiteten

(b) Verteilung der Teilnehmer, welche den 1. Versuch erfolgreich/nicht erfolgreich und mit der Undo-Funktion bearbeiteten

Abbildung 4.20: Erfolgsquote des 1. Versuches mit/ohne Undo-Funktion

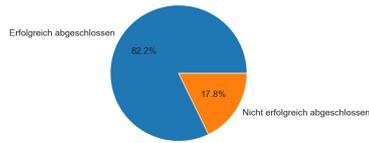
Abbildung 4.21 zeigt die Erfolgsquote für den zweiten Versuch ebenfalls unterteilt in die Nutzer, welche die Undo-Funktion einsetzen (links) und diejenigen, welche sie nicht verwendeten (rechts).

Wie beide Abbildungen zeigen, konnten sowohl die Gruppe mit als auch die Gruppe ohne Undo-Funktion ihre Erfolgsquote im zweiten Durchlauf im Vergleich zum ersten steigern. Die Nutzerperformance verbesserte sich demnach im zweiten Versuch insgesamt gegenüber dem ersten.

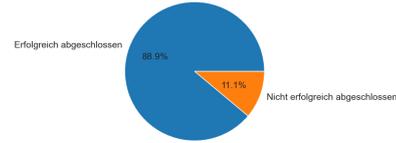
Teilnehmer	Durchführungszeit (%)	Kognitive Belastung (%)	Erlernbarkeit (%)
0	89.00	100.00	100.00
1	88.00	100.00	100.00
2	97.00	60.00	100.00
3	90.00	80.00	80.00
4	95.00	80.00	80.00
5	98.00	40.00	100.00
6	56.00	80.00	80.00
7	64.00	40.00	60.00
8	100.00	60.00	80.00
9	39.00	40.00	60.00
10	51.00	80.00	80.00
11	98.00	60.00	80.00
12	90.00	60.00	60.00
13	73.00	60.00	100.00
14	79.00	100.00	100.00
15	33.00	40.00	60.00
16	69.00	80.00	80.00
17	85.00	60.00	60.00
18	98.00	60.00	80.00
19	99.00	60.00	80.00
20	80.00	40.00	60.00
21	81.00	40.00	60.00
22	80.00	40.00	60.00
23	74.00	40.00	80.00
24	96.00	60.00	60.00
25	82.00	60.00	60.00
26	0.00	20.00	40.00
27	93.00	60.00	100.00
28	74.00	60.00	80.00
29	95.00	60.00	80.00
30	38.00	40.00	40.00
31	61.00	60.00	80.00
32	71.00	60.00	80.00

Tabelle 4.17: Ergebnisse der Durchführungszeiten und Teilnehmerbewertungen zur kognitiven Belastung und Erlernbarkeit (in %)

Prozent Teilnehmer, die den 2. Versuch erfolgreich abgeschlossen haben ohne Undo-Funktion



Prozent Teilnehmer, die den 2. Versuch erfolgreich abgeschlossen haben mit Undo-Funktion



(a) Verteilung der Teilnehmer, welche den 2. Versuch erfolgreich/nicht erfolgreich und ohne Undo-Funktion bearbeiteten

(b) Verteilung der Teilnehmer, welche den 2. Versuch erfolgreich/nicht erfolgreich und mit Undo-Funktion bearbeiteten

Abbildung 4.21: Erfolgsquote des 2. Versuches mit/ohne Undo-Funktion

Vergleichende und kombinierende Metriken

Tabelle 4.17 zeigt die Kombination der Auftragszeit und subjektiven Bewertungen der kognitiven Belastung und Erlernbarkeit (in %) für die erste Versuchsvariante. Je höher der Prozentwert, desto näher liegt die Bearbeitungszeit bei der besten erreichten Auftragszeit. Dasselbe gilt bei den subjektiven Bewertungen. Die Tabelle zeigt teils einen Trend, demnach eine längere Bearbeitungszeit auch zu einer höheren kognitiven Belastung führt. Alle Teilnehmer bis auf zwei Ausnahmen (Teilnehmer 6 und 10), deren Zeit 65% oder weniger an der besten erreichten Zeit liegt, haben die kognitive Belastung unter 50% bewertet. Das bedeutet, diese Teilnehmer haben die kognitive Belastung als hoch empfunden. Generell zeigt die Tabelle, dass dieser Versuch als eher kognitiv belastend empfunden wurde. 25 von 32 Teilnehmern haben für die kognitive Belastung eine Bewertung von 60% oder weniger gegeben. Das heisst, für einen Grossteil der Teilnehmer war der Versuch, unabhängig von der erreichten Zeit, mittel bis stark kognitiv belastend. Dieses Ergebnis wird auch von den Bewertungen zur Erlernbarkeit unterstützt. 12 Teilnehmer bewerteten die Erlernbarkeit der Schnittstelle mit 60% oder weniger.

4.8.4 Interpretation der Ergebnisse

Die Ergebnisse zu den Versuchen um das Prinzip der Kontrolle und Steuerbarkeit sollen nun interpretiert und begründet werden.

Auftragszeit

Die Ergebnisse zeigen deutlich, dass die Teilnehmer, die die Undo-Funktion nutzten, im Durchschnitt kürzere Bearbeitungszeiten aufweisen als jene, die ohne diese Funktion arbeiteten. Ein möglicher Grund dafür liegt im insgesamt geringeren Interaktionsaufwand bei der Nutzung der Undo-Funktion. Wird diese eingesetzt, entfällt der Erkennungs- und Lernprozess alternativer Wege, um die gelöschten Informa-

tionen wiederzufinden – beispielsweise über die Liste der gelöschten E-Mails. Auch werden insgesamt weniger Mausklicks benötigt. Viele der Probanden, die diesen Weg wählten, mussten sich zudem daran erinnern, dass neu gelöschte Elemente stets am Anfang dieser Liste erscheinen. Auch dieser Erinnerungs- und Erkennungsprozess beansprucht Zeit und kognitive Ressourcen.

Im Vergleich dazu erinnern sich die Nutzer beim Löschen vermutlich daran, aus welcher Position die jeweilige E-Mail entfernt wurde. Wird anschliessend die Undo-Funktion verwendet, ist zwar ebenfalls ein Erkennungsprozess erforderlich, um die wiederhergestellte E-Mail zu lokalisieren – der damit verbundene kognitive Aufwand bleibt jedoch vermutlich geringer, da die Teilnehmer wahrscheinlich erwarten (basierend auf ihrem mentalen Modell), dass die E-Mail wieder an der ursprünglichen Position erscheint. Des Weiteren ist in der Abbildung 4.19 zu sehen, dass speziell die Teilnehmenden, welche die Undo-Funktion nicht nutzen, ihre Interaktionsgeschwindigkeit im zweiten Versuch im Vergleich zum ersten Versuch beschleunigen konnten.

Tabelle 4.14 und Tabelle 4.15 bestärken diese Aussage, denn sie zeigen, dass die Interaktionsgeschwindigkeit im zweiten Versuch im Vergleich zum ersten teilweise erhöht werden konnte. Das ist vermutlich, wie bereits erwähnt, auf einen Lernprozess und Erinnerungsprozess zurückzuführen. Es ist plausibel anzunehmen, dass der genannte Einprägungs- und Lernprozess einigen Nutzern erlaubte, den Menü-Ordner im zweiten Versuch schneller zu erkennen und dessen Funktionsweise zu verstehen. Speziell die Personen, welche das Vorgehen über die gelöschten E-Mails wählten, konnten ihre Effizienz steigern. Diese Aussage wird auch von Tabelle 4.16 unterstützt. Denn wahrscheinlich ist es, dass die Teilnehmer sich während des zweiten Versuchs daran erinnerten, wo der Ordner und der Link „gelöschte Elemente“ liegen und wie die Menüstruktur funktioniert. Auch das Suchen der gelöschten E-Mail in der Liste konnte vermutlich effizienter erfolgen, da die Nutzer in der zweiten Versuchsvariante bereits wussten, dass die gelöschte E-Mail am Anfang der Liste zu finden ist. Das Erhöhen der Interaktionsgeschwindigkeit der Teilnehmenden, welche für beide Versuchsvarianten die Undo-Funktion einsetzen, ist vermutlich ebenfalls auf einen Lernprozess zurückzuführen. Auch die farbliche Hervorhebung der wiederhergestellten E-Mail unterstützte vermutlich die Nutzer beim schnelleren Erkennen und Lokalisieren der zielführenden E-Mail. Insgesamt ermöglicht der Einsatz der Undo-Funktion sowie die damit verbundene Kontrolle über Systemhandlungen, effizient zu einem früheren Punkt im Handlungszyklus zurückzukehren, um von dort aus einen neuen Zyklus zu beginnen oder fortzusetzen.

Auftragserfolg

Grafik 4.20 und 4.21 zeigen, dass die Erfolgsquote im zweiten Versuch gegenüber dem ersten Versuch gesteigert werden konnte. Eine Erklärung hierfür lässt sich vermutlich in der unterschiedlichen Funktionsweise der beiden Versuchsschnittstellen und einem Erinnerungsprozess, der durch die Nutzer erfolgte, finden. In beiden Versuchen wurde die neu gelöschte E-Mail, sofern diese nicht über die Undo-Funktion wiederhergestellt wurde, am Anfang der Liste hinzugefügt. Im ersten Versuch wurde die neu gelöschte E-Mail dabei nicht farblich hervorgehoben, in der zweiten Versuchsvariante hingegen schon. Bei einigen Teilnehmern führte diese fehlende farbliche Hervorhebung zu einem Usability-Issue. Diese Teilnehmer waren sich unmittelbar nach dem Aufrufen der gelöschten Elemente im ersten Versuch nicht sicher, wo sich die gelöschte E-Mail in der Liste befindet. Die Teilnehmer versuchten sich danach an Kennzeichnungen der E-Mail zu erinnern und wählten basierend auf diesem Erinnerungsprozess eine Mail aus, welche jedoch in einigen Fällen nicht die korrekte war. Die farbliche Hervorhebung in der zweiten Versuchsvariante lenkte dann die Aufmerksamkeit der Nutzer vermutlich gezielt auf die gelöschte E-Mail. So konnte die korrekte E-Mail entsprechend lokalisiert werden. Die erwähnte farbliche Hervorhebung führte vermutlich auch bei den Nutzern, welche die Undo-Funktion verwendeten, zu einer besseren Erfolgsquote. Insgesamt wurde die Usability im zweiten Versuch im Vergleich zum ersten gezielt durch den Einsatz des Restorff-Effektes erhöht.

Vergleichende und kombinierte Metriken

Die Ergebnisse aus Tabelle 4.17 zeigen möglicherweise die Konsequenzen einerseits aus dem Gesetz von Miller und andererseits die Konsequenzen eines Usability-Issues, welches bei etwa 30% der Teilnehmer, die die Undo-Funktion nicht einsetzen, während des ersten Versuchs in Erscheinung trat. Nach Millers Gesetz werden umso mehr kognitive Ressourcen benötigt, je länger die aktive Aufmerksamkeit der Nutzer gefordert wird und umso mehr Informationen verarbeitet und im Kurzzeitgedächtnis behalten werden müssen.

Usability-Issue

Das erwähnte Usability-Problem zeigte sich wie folgt: Wie bereits im Testaufbau erläutert, sind die Menüeinträge – insbesondere der Eintrag „gelöschte Elemente“ – im Ausgangszustand des Menü-Ordners verborgen. Während des ersten Versuchs stellte sich heraus, dass diese Funktionsweise nicht dem mentalen Modell von etwa 30% der Teilnehmenden entsprach. Wie Tabelle 4.4 zu entnehmen ist, lag die

durchschnittliche Zeit, um den Eintrag „Gelöschte Elemente“ zu finden und anzuklicken, im ersten Versuch vergleichsweise hoch. Dies deutet darauf hin, dass eine Diskrepanz zwischen der vom System vorgesehenen und der von den Nutzenden erwarteten Funktionsweise bestand. Die Funktionsweise des Systems entsprach also nicht dem mentalen Modell der Nutzer. Die tatsächliche Bedienlogik musste also zunächst erlernt werden, was zusätzliche kognitive Ressourcen erforderte.

4.9 Gesetz von Jakob und Mentale Modelle

In diesem Abschnitt geht es um das Gesetz von Jakob sowie um mentale Modelle von Nutzern. In einem ersten Schritt wird durch zwei Versuchsvarianten untersucht, inwiefern sich die Erwartungen an die Systemfunktionalität und das Verhalten der Nutzer abhängig von der Art und Komplexität des zu lösenden Problems anpassen und ändern.

Im einem zweiten Teil werden verschiedene Umsetzungen von Webschnittstellen untersucht. Dabei werden unterschiedliche Konzepte zur Informations- und Navigationsstruktur sowie verschiedene Layouts miteinander verglichen. Ziel ist es herauszufinden, welche dieser Varianten von den Benutzern am schnellsten verstanden und am effizientesten genutzt werden können. Ausserdem wird analysiert, welche Gestaltungsansätze es den Anwendern ermöglichen, auf ihr vorhandenes Wissen aus früheren Erfahrungen zurückzugreifen. Insgesamt soll ermittelt werden, welche Schnittstellenmuster am ehesten den mentalen Modellen der Nutzer entsprechen. Der Vergleich der Performance während der Nutzung von unterschiedlichen Strukturen und Layouts erlaubt es sodann, mögliche Aussagen über die mentalen Modelle der Nutzer und der Usability der jeweiligen Schnittstellen zu treffen [30].

Ergänzend sollen die Experimente auch dazu beitragen, Erkenntnisse darüber zu gewinnen, ob Informationsarchitekturen, die weniger stark verbreiteten Konventionen folgen, durch den gezielten Einsatz von Gestaltungselementen wie ausgeprägtem Angebotscharakter (Affordances) dennoch leicht verständlich und nutzbar gestaltet werden können.

4.9.1 Testaufbau

Benutzerinitiierte Nutzerkonfigurationen

Mentale Modelle sind eng mit den Erwartungen, welche Nutzer an ein System haben, verknüpft. Um die Erwartungen zu untersuchen, wurden zwei experimentelle Tests entwickelt, in denen die Teilnehmenden eigenständig eine Webschnittstelle eines Webshops konfigurieren müssen. Dabei wählen sie aus verschiedenen Typen

von Interaktionselementen – wie Menü- und Navigationskomponenten – und legen deren Anordnung innerhalb des Webshops fest. Darüber hinaus können zusätzliche Funktionalitäten wie Such- oder Filterkomponenten hinzugefügt werden. Der Auswahlprozess sowie die zur Verfügung stehenden Komponenten sind in Anhang D zum Thema mentale Modelle veranschaulicht.

Das Ziel beider Testvarianten ist identisch: Die Teilnehmenden sollen einen eigenen Online-Webshop konfigurieren und diesen im Anschluss nutzen, um ein bestimmtes Produkt zu finden und in den Warenkorb zu legen. Den Teilnehmenden wird ein Konfigurationsmenü präsentiert, wo die Konfiguration erfolgt. Anschließend bestätigen diese die Konfiguration und der Webshop wird gemäss ihrer Konfiguration vom System aufgebaut. Nun kann der Web-Shop eingesetzt werden, um das verlangte Produkt zu suchen und in den Warenkorb zu legen. Dabei reicht es nicht aus, ein beliebiges Produkt auszuwählen – das gesuchte Produkt muss vorgegebene Spezifikationen erfüllen. Die Anzahl und Art dieser Spezifikationen unterscheiden sich in den beiden Testvarianten. In der zweiten Versuchsvariante ist die Anzahl der Spezifikationen höher als in der ersten Variante. Den Teilnehmenden ist es während der Produktsuche jederzeit möglich, zurück zum Konfigurationsmenü zurückzukehren, um weitere Elemente der Schnittstelle hinzuzufügen.

Im ersten Test muss das Produkt folgende Kriterien erfüllen:

- Type: Notebook
- Marke: Apple

Im zweiten Test wurden die Anforderungen bewusst erweitert, um die Komplexität der Aufgabe zu steigern:

- Type: Notebook
- Marke: Apple
- Arbeitsspeicher: mind. 16 GB
- Speicherkapazität: mind. 256 GB

Durch die erhöhte Anzahl an Spezifikationen im zweiten Test soll untersucht werden, inwiefern sich die Konfigurationsstrategie der Versuchsteilnehmer sowie das Suchverhalten der Teilnehmenden in Abhängigkeit von der Aufgabenkomplexität verändert.

Vergleich von 5 Webschnittstellen

Um die mentalen Modelle der Probanden weiter zu untersuchen, wurden fünf unterschiedliche Varianten eines Online-Webshops implementiert. Diese Webshop-Varianten unterscheiden sich hinsichtlich ihrer Navigationslogik sowie in weiteren funktionalen Aspekten und werden im Folgenden näher beschrieben. Alle Varianten sind in Anhang D, Abschnitt D.8 abgebildet.

Die erste Variante orientiert sich an klassischen Gestaltungsprinzipien für Webschnittstellen. Die grafische Oberfläche verfügt am oberen Bildschirmbereich über einen horizontalen Navigationsheader, der eine Suchfunktion sowie ein Warenkorb-Icon in der rechten oberen Ecke enthält. Die Suchfunktion ist mit einer Autovervollständigung ausgestattet, sodass bei der Eingabe von Suchbegriffen passende Produktvorschläge angezeigt werden. Zusätzlich enthält die Schnittstelle ein textbasiertes Off-Canvas-Menü [2] am linken Rand des Webshops. Dieses Gestaltungsprinzip sieht vor, dass die Navigationselemente ausserhalb des sichtbaren Inhaltsbereichs platziert und erst über ein Menü-Icon eingeblendet werden. Im Ausgangszustand bleiben die Navigationslinks verborgen; wird das Menü geöffnet, verschiebt sich der Inhaltsbereich leicht nach rechts. Ein erneutes Betätigen des Icons blendet die Navigationselemente wieder aus. Die Navigation selbst folgt einer hierarchischen Struktur: Bei Auswahl einer Kategorie werden die aktuell sichtbaren Menüeinträge durch die zugehörigen Unterkategorien ersetzt, wobei die gewählte Kategorie farblich hervorgehoben wird. Die übergeordnete Kategorie (Parent-Link) bleibt dabei stets an oberster Stelle im Menü sichtbar. Wählt der Nutzer eine Produktkategorie, erscheinen zudem passende Filteroptionen, mit denen sich die Produktanzeige gezielt eingrenzen lässt. Bei Betätigung des Warenkorb-Icons wird der Warenkorb geöffnet.

Die zweite Webshop-Variante verfügt ebenfalls über einen Navigationsbalken mit Webshop-Titel, Suchfunktion und einem Warenkorb-Icon links neben dem Suchfeld. Im Gegensatz zur ersten Variante bietet diese Suchfunktion jedoch keine Autovervollständigung. Statt eines Off-Canvas-Menüs wird hier ein sogenanntes Mega-Dropdown-Menü verwendet. Dieses Menü weist ebenfalls eine mehrstufige Struktur auf, visualisiert jedoch bereits im Ausgangszustand sämtliche Hauptkategorien. Bei Mouseover über einen Hauptmenüpunkt öffnet sich das Menü und zeigt die zugehörigen Unterkategorien vertikal an. Ein Klick auf eine dieser Kategorien führt zur Anzeige der entsprechenden Unterkategorien. Verlässt der Mauszeiger den Menübereich, schliesst sich das Menü automatisch. Wird hingegen der Mauszeiger zwischen verschiedenen Hauptkategorien bewegt, bleibt das Menü geöffnet und aktualisiert dynamisch die angezeigten Inhalte. Auch in dieser Variante werden

Filteroptionen eingeblendet, sobald eine Produktkategorie ausgewählt wurde.

Die dritte Webshop-Variante unterscheidet sich von der ersten ausschliesslich durch die Menükomponente sowie die Suchfunktion. Die Suchfunktion verzichtet hier auf eine Autovervollständigung. Das Menü ist fest eingeblendet und benötigt kein Icon zum Öffnen oder Schliessen – im Ausgangszustand werden sämtliche Hauptkategorien als anklickbare Links untereinander sichtbar dargestellt.

Die vierte Variante besitzt ebenfalls einen Navigationsbalken mit einer Suchfunktion (inklusive Autovervollständigung) sowie dem Webshop-Titel. Das Icon für den Warenkorb ist links neben dem Suchfeld platziert. Die Menükomponente entspricht derjenigen der dritten Variante. Im Gegensatz zu den bisher beschriebenen Varianten befindet sich die Menükomponente hier jedoch am rechten Rand der Benutzeroberfläche, während der Navigationsbalken mitsamt Suchfunktion am unteren Bildschirmrand positioniert ist. Die fünfte Variante des Webshops entspricht in Aufbau und Funktionalität der zuvor beschriebenen vierten Version. Der einzige Unterschied besteht darin, dass es sich bei der Menükomponente um ein Off-Canvas-Menü wie in der ersten Variante handelt.

Auf Grundlage dieser fünf Varianten wurden ebenso fünf Versuchsexperimente konzipiert. Ziel aller Versuche ist es, ein Produkt mit bestimmten, vorgegebenen Eigenschaften zu finden und in den Warenkorb zu legen. Welche Eigenschaften dies im Einzelnen sind, wird jeweils in der Versuchsanleitung definiert. Die Anzahl der geforderten Produktspezifikationen steigt von Versuch zu Versuch schrittweise an: Im ersten Experiment sind zwei Spezifikationen vorgegeben, im zweiten drei, im dritten vier und im vierten Versuch fünf. Im fünften und letzten Experiment wird die Anzahl der Spezifikationen wieder auf zwei reduziert. Diese Variation der Spezifikationsanzahl dient der gezielten Anpassung der Aufgabenkomplexität. Dadurch soll erfasst werden, wie sich das Suchverhalten der Teilnehmenden in Abhängigkeit von der Komplexität der gestellten Aufgabe verändert.

4.9.2 Usability- und User Experience-Metriken

Issue-bezogene Metriken

Für die Analyse der Nutzerkonfigurationen wird zum einen erfasst, wie häufig bestimmte Elemente von den Teilnehmenden gewählt und der Schnittstelle hinzugefügt wurden. Zudem wird gemessen, wie oft diese Elemente tatsächlich eingesetzt wurden, um die gestellte Aufgabe zu bewältigen. Diese beiden Kennzahlen, so-

genannte Issue-bezogene Metriken, werden miteinander verglichen, um potenzielle Diskrepanzen zwischen dem wahrgenommenen und dem tatsächlichen Nutzen der Elemente zu identifizieren. Eine Verhaltensanalyse der Nutzer kann ebenfalls helfen, Aussagen über mentale Modelle zu treffen. Hier bietet es sich an, das Suchverhalten der Nutzer während der fünf Versuchsvarianten zu analysieren und zu vergleichen, um etwaige Muster im Handeln der Nutzer zu erkennen. Konkret soll hier untersucht werden, wie sich die Nutzung von Suchfunktionen, Filtern und die Benutzung der textbasierten Menükomponenten mit steigender Aufgabenkomplexität ändern. Diese Arten von Metriken können ebenfalls der Kategorie Issues-based-Metriken zugeordnet werden.

Zum anderen erscheint es sinnvoll, auch Metriken zu Usability-Issues in die Analyse einzubeziehen. Usability-Issues können auf Lücken zwischen den mentalen Modellen der Systementwickler und denen der Nutzer hinweisen. Die Identifikation solcher Diskrepanzen ermöglicht eine gezielte Anpassung des Designs an die mentalen Modelle der tatsächlichen Zielgruppe. In diesem Zusammenhang ist es sinnvoll zu bestimmen, wie viele eindeutig identifizierbare Usability-Issues auftreten und bei wie vielen Testpersonen ein bestimmtes Usability-Issue beobachtet werden kann. Um die fünf zuvor beschriebenen Schnittstellen hinsichtlich ihrer Übereinstimmung mit den mentalen Modellen der Nutzer zu bewerten, ist eine ganzheitliche Betrachtung der User Experience erforderlich. Aussagen über die zugrunde liegenden mentalen Modelle setzen eine umfassende Analyse voraus, die verschiedene Aspekte der User Experience einbezieht. Aus diesem Grund scheint es sinnvoll, kombinierte und vergleichende Metriken ebenfalls heranzuziehen. Insbesondere könnte eine Kombination von Performance- und Erlernbarkeitsmetriken mit selbstberichteten Metriken (Self-reported-Metriken) hilfreiche Erkenntnisse liefern.

Performance-Metriken

Als Performance-Metriken werden dabei die Auftragsbearbeitungszeit, die Anzahl der Fehlklicks sowie die Gesamtanzahl der Mausklicks berücksichtigt. Zur Messung der Erlernbarkeit dient die Zeitspanne bis zum ersten Mausklick auf eine Schaltfläche, die einen Handlungszyklus initiiert. Grundsätzlich lassen sich hier auch mehr Erlernbarkeitsmetriken entwickeln, beispielsweise durch die Messung wie sich ein investierter Aufwand oder Verhalten (Anzahl Fehlklicks, Anzahl Mausklicks, subjektive Bewertungen usw.) im Verlauf der fünf Versuche ändert.

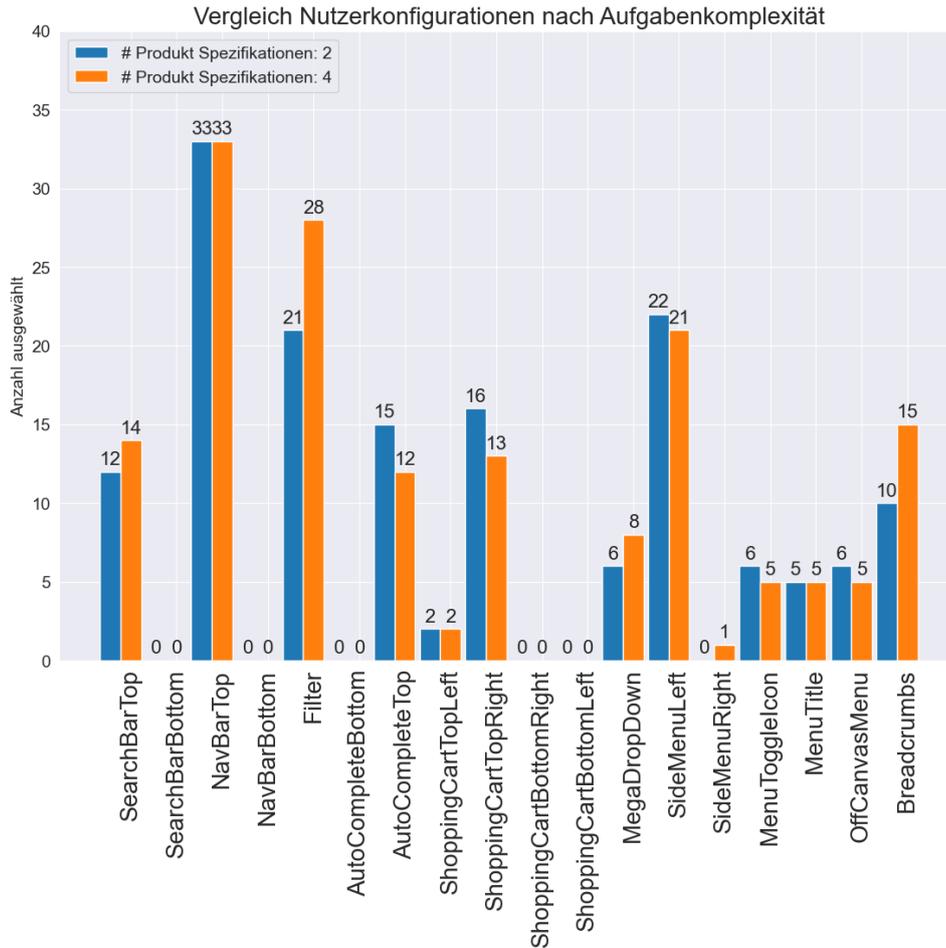


Abbildung 4.22: Häufigkeit der konfigurierten Schnittstellen-Elemente nach Aufgabenkomplexität (Anzahl Produktspezifikationen)

4.9.3 Ergebnisse

Nutzerkonfigurationen

Als erstes werden die Resultate der Nutzerkonfigurationen präsentiert. Diese Ergebnisse sind in Abbildung 4.22 ersichtlich.

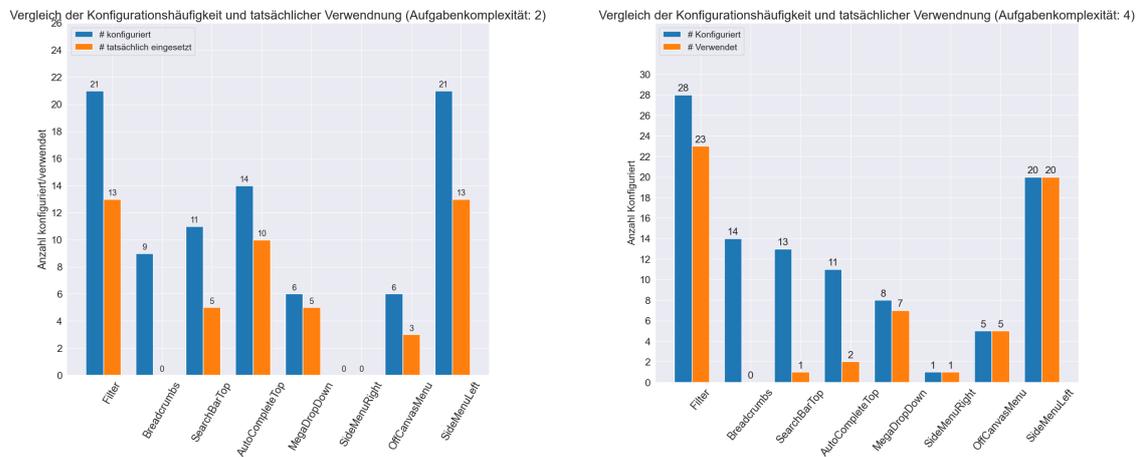
Die Grafik vergleicht die Konfiguration der Schnittstellen in Abhängigkeit von der Aufgabenkomplexität, die durch die Anzahl der Produktspezifikationen definiert ist, welche das zu findende Produkt erfüllen muss. Die X-Achse zeigt die konfigurierbaren Elemente und die Y-Achse, wie oft ein Element konfiguriert wurde. Auffällig ist, dass die Gesamtanzahl der verwendeten Elemente in beiden Testvarianten vergleichbar ausfällt. So wurde beispielsweise eine Suchfunktion in der ersten Variante zwölfmal und in der zweiten Variante vierzehnmal genutzt. Ein ähnliches Muster zeigt sich bei den textbasierten Menükomponenten (SideMenuLeft, MegaDropDown, Off-Canvas-Menü) sowie beim Warenkorb und der Navigationsleiste

(NavBarTop). Die deutlichste Abweichung ist bei der Nutzung der Filterfunktionen erkennbar: In der ersten Testvariante mit einer Aufgabenkomplexität von zwei wurden die Filter insgesamt 21-mal konfiguriert, während in der zweiten Variante mit einer höheren Aufgabenkomplexität von vier die Filterfunktionen 28-mal konfiguriert wurden. Dabei ist zu beachten, dass Abbildung 4.22 ausschliesslich die Häufigkeit der vorgenommenen Konfigurationen darstellt, nicht jedoch die tatsächliche Nutzung der jeweiligen Elemente. Im nächsten Schritt wird daher das Verhältnis zwischen der Konfigurationshäufigkeit und der tatsächlichen Nutzung der einzelnen Elemente näher analysiert.

Abbildung 4.23 vergleicht für beide Konfigurations- und Suchaufträge mit unterschiedlicher Aufgabenkomplexität (zwei und vier) für bestimmte Elemente, wie oft diese ausgewählt und tatsächlich verwendet wurden. Wie der Abbildung 4.23a entnommen werden kann, zeigen sich bereits bei einer Aufgabenkomplexität von zwei, also zwei Produktspezifikationen, einige Diskrepanzen in der Konfiguration und der tatsächlichen Verwendung von Elementen. Was hier ins Auge fällt, sind die Unterschiede in der Verwendung und Konfiguration von Filterfunktionen, Suchfunktionen (SearchBarTop, AutoCompleteTop), Breadcrumbs (Navigationslinks) und textbasierten Menüs, wie dem Off-Canvas-Menü oder einem links platzierten Seitenmenü (SideMenuLeft). Die Breadcrumbs, das sind Navigationslinks, wurden zum Beispiel neun Mal gewählt, jedoch kein einziges Mal verwendet. Filterfunktionen wurden 21-mal gewählt, doch nur 13-mal eingesetzt. Auch ein textbasiertes Navigationsmenü wie das linke Seitenmenü (SideMenuLeft) wurde 21-mal konfiguriert, doch nur 13-mal verwendet.

Werden die beiden Abbildungen 4.23a und 4.23b miteinander verglichen, so fallen einige Tatsachen auf: Die Diskrepanzen in der Konfiguration und der tatsächlichen Verwendung ausgewählter GUI-Elemente fallen unterschiedlich für die beiden Aufgaben mit verschiedener Aufgabenkomplexität aus. Im zweiten Versuchstest mit Aufgabenkomplexität vier wurde eine Suchfunktion (SearchBarTop oder AutoCompleteTop) ähnlich oft (24-mal) konfiguriert wie im ersten Versuchstest (25-mal) mit Aufgabenkomplexität zwei. Doch tatsächlich verwendet wurde eine Suchfunktion im zweiten Versuch nur noch drei Mal. Eine Suchfunktion wurde also im Versuch mit geringerer Komplexität viel häufiger eingesetzt als im Versuch mit höherer Komplexität. Auch die Konfiguration und Verwendung von textbasierten Menüs resultiert in der zweiten Versuchsvariante unterschiedlich aus. Im ersten Versuch wurde ein textbasiertes Menü (Off-Canvas, Mega-Drop-Down oder SideMenuLeft) insgesamt 33-mal ausgewählt und 21-mal verwendet. Im zweiten Versuch mit höherer Aufgabenkomplexität wurde ein textbasiertes Menü 34-mal ausgewählt, doch insgesamt

32-mal verwendet. Auch Filterfunktionen wurden in der zweiten Versuchsvariante mehr konfiguriert (28-mal) als tatsächlich eingesetzt (23-mal) als im Vergleich zum ersten Versuch, wo die Filterfunktionen 21-mal konfiguriert wurden, jedoch nur 13-mal eingesetzt wurden.



(a) Vergleich Konfigurationshäufigkeit und Verwendungshäufigkeit der konfigurierten Elemente mit Aufgabenkomplexität zwei

(b) Vergleich Konfigurationshäufigkeit und Verwendungshäufigkeit der konfigurierten Elemente mit Aufgabenkomplexität vier

Abbildung 4.23: Vergleich Konfigurationshäufigkeit und Verwendungshäufigkeit der konfigurierten Elemente nach Aufgabenkomplexität

Usability-Issue

Während der Durchführung der beiden Konfigurationsaufgaben mit anschliessender Suche eines Produktes mit der konfigurierten Schnittstelle konnte ein Usability-Issue beobachtet werden, welches bei etwa 45% der Versuchsteilnehmenden, welche ein textbasiertes Menü verwendeten, aufgetreten ist. Das Usability-Issue wurde wie folgt beobachtet: Wie die Grafik 4.22 zeigt, wurden die Filterfunktionen in beiden Versuchsvarianten oft konfiguriert. Nachdem die Probanden ihre Schnittstellenkonfiguration abgeschlossen haben und mit der Produktsuche über die konfigurierte Schnittstelle begannen, wollten sie direkt die konfigurierten Filterfunktionen einsetzen, um das gewünschte Produkt zu filtern und auszuwählen. Die Nutzer haben also erwartet, die Suche direkt mit den hinzugefügten Filtern beginnen zu können. Das System erlaubte diese Interaktion jedoch nicht direkt. Denn unmittelbar nachdem die Konfiguration abgeschlossen wurde, präsentierte der konfigurierte Web-Shop die Hauptseite. Dort wurden standardmässig noch keine Filter visualisiert, selbst wenn diese der Schnittstelle zuvor hinzugefügt wurden. Damit das System die Filter visualisiert, muss erst über das textbasierte Menü eine Produktkategorie ausgewählt werden. Erst dann werden die Filter präsentiert (vgl. Anhang D, Jakob's Gesetz,

Abbildungen D.25 und D.26).

Bei etwa 45% der Teilnehmer, welche ein textbasiertes Menü einsetzten, führte dieses Systemverhalten zu Verwirrung. Sie entschieden, zurück zur Konfiguration zu navigieren, um zu verifizieren, ob die Filter tatsächlich hinzugefügt wurden, um dann ihre Konfiguration erneut zu bestätigen und mit der Suche zu beginnen. Bei einigen Teilnehmern erfolgte dieser Handlungszyklus sogar zwei bis drei Mal, bis sie erkannten, dass ein Menüeintrag betätigt werden muss.

Vergleich von 5 Webschnittstellen

Nachfolgend werden die Ergebnisse der fünf Versuchsvarianten, in denen es darum ging, fünf Schnittstellen zu vergleichen und zu evaluieren, welche dieser Schnittstellen am ehesten den mentalen Modellen der Nutzer entsprechen, präsentiert.

Performance-Metriken

Als erstes werden die Auftragszeiten analysiert. Abbildung 4.24 präsentiert die mittleren Bearbeitungszeiten der Aufträge nach Schnittstellentyp. Wie zu erkennen ist, steigen die Auftragszeiten von links nach rechts bis zur vierten Schnittstelle an. Anschliessend im letzten Versuch, unter Verwendung der Off-Canvas-Menüstruktur, sinkt die Auftragszeit wieder. Als Vorgriff in die Ergebnisinterpretation kann dieses Muster damit erklärt werden, dass die Aufgabenkomplexität von links nach rechts bis zur vierten Schnittstelle schrittweise erhöht wird. Der fünfte Versuch weist dann wieder dieselbe Aufgabenkomplexität wie der erste Versuch auf. Die benötigte Bearbeitungszeit steigt und sinkt (wie erwartet) mit der Aufgabenkomplexität. Tabelle 4.18 zeigt ergänzend, welcher Schnittstellentyp wie viele Nutzer aufweist, deren Bearbeitungszeit über einer Minute liegt.

Schnittstellennr.	Anzahl Teilnehmer mit Bearbeitungszeit > 60 Sekunden
1	7
2	7
3	14
4	19
5	4

Tabelle 4.18: Anzahl Nutzer, die mehr als eine Minute benötigten, die Aufgabe abzuschliessen nach Schnittstellenummer

Abbildung 4.25 vergleicht unterschiedliche Performance-Metriken (Erlernbarkeit, Fehler und Effizienz) nach Schnittstellenvariante miteinander. Eine Metrik für die Erlernbarkeit ist die Anzahl Sekunden, bis die erste Schaltfläche betätigt worden ist, welche einen Suchprozess des Nutzers initiiert. Als Effizienzmetrik dienen die

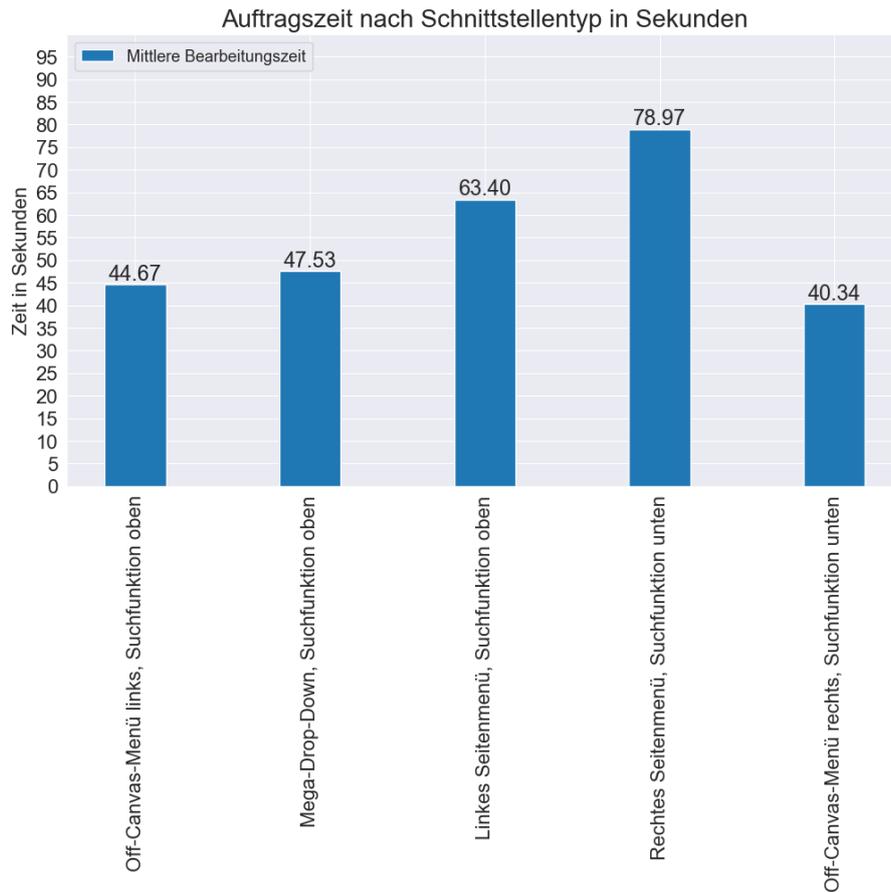


Abbildung 4.24: Auftragszeiten nach Schnittstellentyp

getätigten Mausklicks und als Fehlermetrik wurde die Anzahl Fehlklicks gemessen. Alle Teilnehmer haben die fünf Versuche in der Reihenfolge von links nach rechts wie in der Grafik dargestellt durchgeführt.

Wie der Abbildung zu entnehmen ist, zeigen der erste und der letzte Versuch eine vergleichbare Anzahl Sekunden bis zum ersten Mausklick (17,8 bzw. 16,8 Sekunden im Schnitt). Am längsten benötigten die Nutzer für den Beginn des Suchprozesses bei der Schnittstelle mit dem Mega-Drop-Down-Menü (24,38 Sekunden). Diese wurde jedoch insgesamt am effizientesten genutzt, gemessen an der durchschnittlich geringsten Anzahl an Mausklicks (sieben Klicks). Die geringste Effizienz zeigte sich hingegen bei der vierten Schnittstelle mit einem textbasierten Menü am rechten Rand, bei der im Mittel 12 Mausklicks erforderlich waren. Darüber hinaus weist die Abbildung darauf hin, dass bei der letzten getesteten Schnittstelle – einem Off-Canvas-Menü am rechten Rand – im Durchschnitt die meisten Fehlklicks auftraten.

Abbildung 4.25 lässt vermuten, dass die vierte Schnittstelle die am schwierigsten erlernbare und die am wenigsten an das mentale Modell der Nutzer angepasste Schnittstelle ist. Diese Vermutung lässt sich durch die subjektiven Bewertungen der Studienteilnehmenden mithilfe der Abbildungen 4.26 und 4.27 verifizieren.

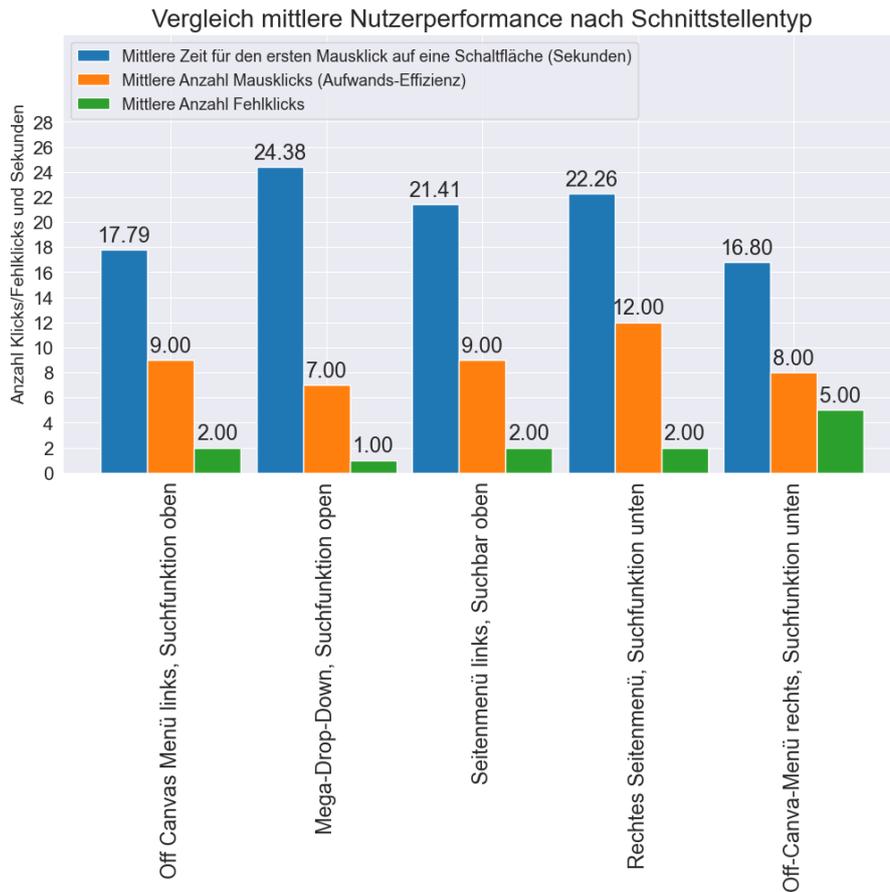


Abbildung 4.25: Vergleich verschiedener Aspekte der Nutzerperformance nach Schnittstellentyp

Sowohl Grafik 4.26 und 4.27 unterstützen die in Abbildung 4.25 dargestellten Ergebnisse. Generell zeigt sich durch die Bewertungen der Nutzer ein Trend, nach dem Schnittstellen mit Menükomponenten am rechten Rand und Suchfunktionen am unteren Bildschirmbereich als weniger benutzerfreundlich empfunden werden.

In den beiden Abbildungen 4.23a und 4.23b wird das bereits beschriebene Muster dargestellt, das sich während der Konfigurationstests zeigte. Es soll untersucht werden, ob sich dieses Verhalten auch während der fünf Tests mit unterschiedlicher Schnittstellengestaltung und Aufgabenkomplexität beobachten lässt. Das erwähnte Muster bezieht sich vor allem auf die Nutzung von Filterfunktionen, textbasierten Menüs und Suchfunktionen in Abhängigkeit von der Aufgabenkomplexität. Weiter oben konnte beobachtet werden, dass der erste Konfigurationsversuch mit geringerer Aufgabenkomplexität die grössere Anzahl an Teilnehmern aufwies, die eine Suchfunktion verwendeten. Im zweiten Versuch mit höherer Komplexität nahm diese Zahl ab, und mehr Probanden griffen auf Filterfunktionen und textbasierte Menükomponenten zurück.

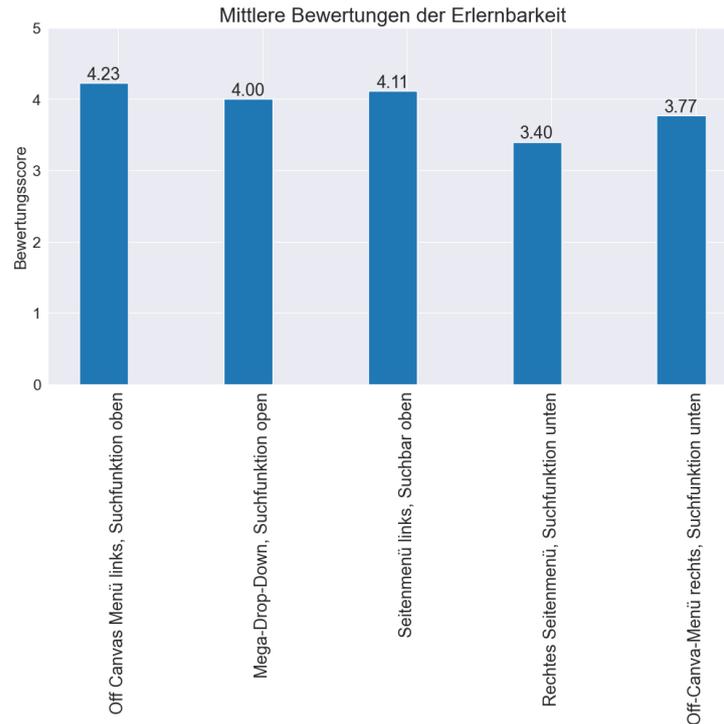


Abbildung 4.26: Nutzerbewertungen der Erlernbarkeit nach Schnittstellentyp

Wie Abbildung 4.28 zeigt, findet sich dieses Muster auch in dieser Experimentreihe. Auf der X-Achse sind die unterschiedlichen Versuchsvarianten (T) mit den unterschiedlichen Aufgabenkomplexitäten (C) angegeben. Die Y-Achse zeigt die jeweilige Anzahl Nutzer, welche eine Suchfunktion (Abbildung 4.28a) oder Filterfunktionen (Abbildung 4.28b) einsetzten. Je höher die Komplexität wird, umso weniger Nutzer verwendeten eine Suchfunktion. Und umso höher die Komplexität wurde, desto mehr Teilnehmer nutzten Filterfunktionen in Kombination mit einem textbasierten Menü. Interessant sind die Messwerte zu den beiden Tests $T1$ und $T2$. Obwohl die Aufgabenkomplexität in $T5$ gleich ist wie bei $T1$, nutzten im letzten Versuch weniger Teilnehmer eine Suchfunktion als beim ersten Versuch.

Vergleichende und kombinierte Metriken

Tabelle 4.19 zeigt die Ergebnisse des vierten Versuchstests, bei dem die Aufgabe die höchste Komplexität (Stufe 5) aufweist. Dargestellt sind die Bearbeitungszeiten sowie die subjektiven Bewertungen der Teilnehmenden hinsichtlich kognitiver Belastung, Erlernbarkeit und der wahrgenommenen Angepasstheit an das eigene mentale Modell. Die Werte in der Tabelle wurden gemäss Gleichung 3.1 berechnet: höhere Prozentwerte entsprechen dabei einer stärkeren Ausprägung im Sinne eines optimalen Werts. Die Daten lassen einen Trend erkennen: Teilnehmende, welche die kognitive Belastung als hoch oder die Erlernbarkeit als gering einschätzten (d.h.

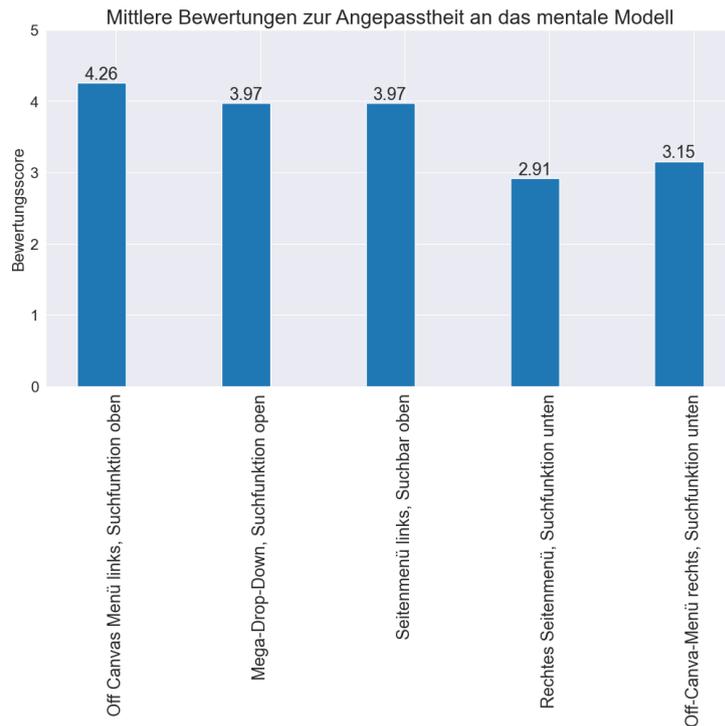


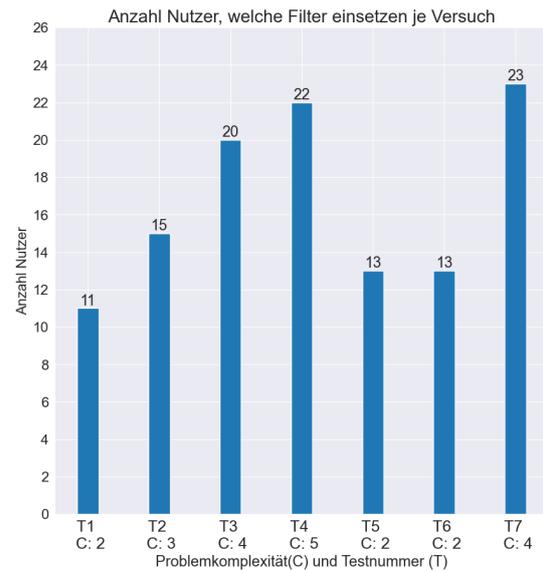
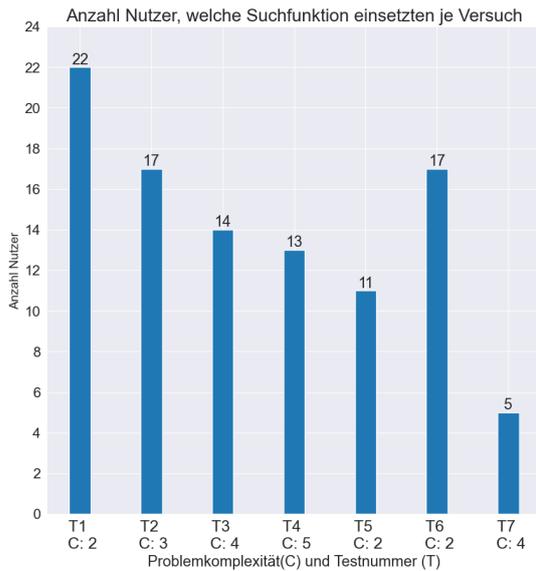
Abbildung 4.27: Nutzerbewertungen der Angepasstheit an das mentale Modell nach Schnittstellentyp

Erlernbarkeit $\leq 60\%$ oder *kognitiveBelastung* $\leq 60\%$), bewerteten tendenziell auch die Übereinstimmung der Schnittstelle mit ihrem mentalen Modell als gering (*mentalesModell* $\leq 60\%$). Diese Beobachtung war zu erwarten: Je besser das Design einer Schnittstelle an das mentale Modell der Nutzer angepasst ist, desto leichter fällt das Erlernen – und entsprechend geringer ist die empfundene kognitive Belastung.

4.9.4 Interpretation der Ergebnisse

Auftragszeit

In den Ergebnissen wurde erwähnt, dass das in Abbildung 4.24 dargestellte Zeitmuster auf die steigende Aufgabenkomplexität zurückzuführen ist. Es erscheint jedoch ebenso plausibel, dass auch die jeweiligen Strukturen der Schnittstellen zu diesen Unterschieden beigetragen haben. So könnte es sein, dass einige Nutzerinnen und Nutzer mehr Zeit benötigten, um mit Schnittstellen zurechtzukommen, die weniger verbreitete Designkonventionen verwenden – wie beispielsweise die vierte Schnittstelle in Abbildung 4.26. Diese Hypothese wird durch Tabelle 4.18 gestützt: Dort ist deutlich zu erkennen, dass die vierte Schnittstelle die höchste Anzahl an Teilnehmenden aufweist, deren Bearbeitungszeit über einer Minute lag. Abbildung 4.29



(a) Anzahl Nutzer, welche eine Suchfunktion einsetzten

(b) Anzahl Nutzer, welche ein textbasiertes Menü und Filterfunktionen einsetzten

Abbildung 4.28: Vergleich Anzahl Nutzer, die ein textbasiertes Menü und Filter oder eine Suchfunktion einsetzen nach Versuchsvariante

verdeutlicht zusätzlich, dass die vierte Schnittstelle im Vergleich zu den übrigen häufiger Bearbeitungszeiten hervorgebracht hat, die oberhalb des sogenannten Interquartilsbereichs (d.h. ausserhalb der mittleren 50% aller Werte) liegen.

Nutzerkonfigurationen

Zunächst sollen die Ergebnisse der Nutzerkonfigurationen interpretiert werden. Ein erneuter Blick auf Abbildung 4.22 legt die Vermutung nahe, dass die funktionalen Erwartungen der Nutzenden an die Schnittstelle in beiden Versuchsvarianten weitgehend übereinstimmen. Viele Elemente wurden in beiden Varianten in ähnlicher Häufigkeit konfiguriert.

Eine auffällige Ausnahme bildet die Nutzung der Filterfunktionen. Hier zeigen sich Unterschiede in Abhängigkeit von der Aufgabenkomplexität. Je höher die Komplexität der Aufgabe – also je mehr Produktspezifikationen im Kurzzeitgedächtnis behalten werden müssen – desto ausgeprägter scheint der Wunsch der Nutzenden zu sein, Objekte zu filtern, um ausschliesslich relevante Produkte angezeigt zu bekommen, um die bereits höhere kognitive Belastung aufgrund der Aufgabenkomplexität reduzieren zu können.

Abgesehen davon lässt sich das in Abbildung 4.22 dargestellte Ergebnis wie folgt interpretieren: Es ist naheliegend anzunehmen, dass viele Teilnehmende bereits über ein mentales Modell eines Webshops verfügen. Dieses mentale Modell basiert auf früheren Nutzungserfahrungen mit vergleichbaren Online-Shops und umfasst Vor-

Nutzer	Durchführungszeit (%)	Kognitive Belastung (%)	Erlernbarkeit (%)	Mentales Modell (%)
1	94	80	40	40
2	94	100	100	100
3	92	60	60	80
4	91	60	80	60
5	81	60	40	40
6	81	60	60	80
7	96	60	40	40
8	94	80	80	60
9	88	80	80	80
10	92	60	60	40
11	89	80	60	40
12	0	40	60	60
13	90	60	100	80
14	88	100	80	80
15	85	80	60	40
16	73	80	80	40
17	51	80	80	80
18	93	80	80	80
19	87	80	60	60
20	70	60	80	60
21	81	40	60	40
22	91	80	100	60
23	69	40	60	20
24	82	40	40	40
25	81	60	60	60
26	83	60	60	40
27	21	60	60	60
28	98	20	100	100
29	100	80	80	80
30	93	60	60	40
31	91	60	40	40
32	96	60	60	60
33	64	60	60	40
34	88	80	80	40
35	82	60	80	80

Tabelle 4.19: Vergleichende und kombinierte Metriken; Durchführungszeiten und Bewertungen der kognitiven Belastung, Erlernbarkeit und zur Angepasstheit an das mentale Modell in % für den vierten Versuchstest

stellungen darüber, welche Funktionen ein Webshop bereitstellen sollte und wie dieser grundsätzlich aufgebaut ist und aussehen sollte. Daher ist es plausibel anzunehmen, dass die Mehrheit der Nutzenden in beiden Versuchsvarianten ihnen bekannte und weit verbreitete Navigationskomponenten wie Menüs, Breadcrumbs, Navigationsleisten oder Suchfunktionen integriert hat, um die Oberfläche entsprechend ihrem mentalen Modell eines funktionalen Webshops zu gestalten. Die Abbildungen 4.23a und 4.23b verdeutlichen jedoch, dass Diskrepanzen zwischen den Erwartungen an die äussere Gestaltung einer Schnittstelle (abhängig vom Nutzungskontext, z.B. Online-Shopping) und den Erwartungen an deren tatsächliche Funktionalität bestehen. Es ist daher möglich anzunehmen, dass der konkrete Gebrauch von Funktionalitäten nicht nur vom allgemeinen Nutzungskontext, sondern auch von der spezifischen Aufgabe – etwa der Art des zu kaufenden Produkts – beeinflusst wird.

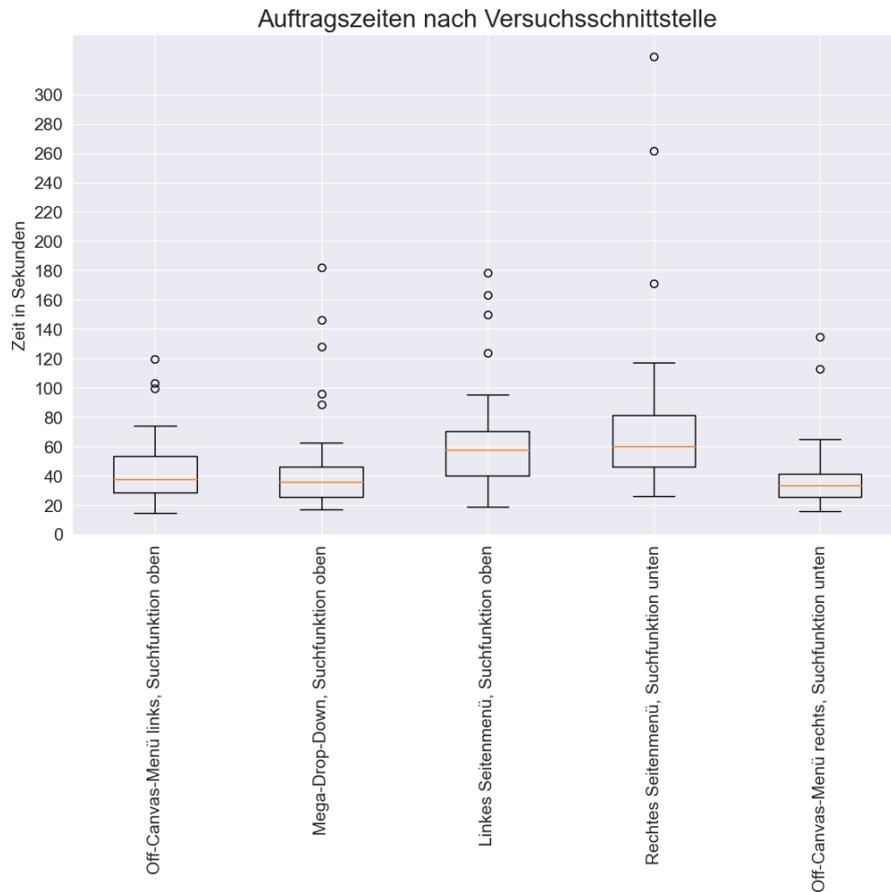


Abbildung 4.29: Auftragszeiten nach Versuchsschnittstelle

Für Designer ergibt sich daraus die Anforderung, die Komplexität eines Systems oder Produkts so gering wie möglich zu halten, um die Usability zu erhöhen. Ziel sollte es sein, ausschliesslich die Funktionalitäten bereitzustellen, die von der Zielgruppe tatsächlich benötigt werden. Nach dem Gesetz von Tesler muss jedoch berücksichtigt werden, dass eine minimale Systemkomplexität bestehen bleiben muss. Funktionalitäten können also nicht beliebig reduziert werden. Unter Einbeziehung mentaler Modelle der Nutzenden ist daher ein Mittelweg zu finden: Einerseits soll die Komplexität im Sinne Teslers reduziert werden, andererseits muss die Gestaltung der Schnittstelle mit den Erwartungen und Erfahrungswerten der Nutzenden übereinstimmen, damit diese ihr mentales Modell anwenden können und die Interaktion effizient und intuitiv bleibt.

Usability-Issue

Das in den Ergebnissen beschriebene Usability-Issue ist ein Beispiel einer Lücke zwischen dem mentalen Modell der Nutzer und dem mentalen Modell der Systementwickler. Denn die Erwartungen der Nutzer, wie das System funktionieren sollte,

unterschieden sich von der Art und Weise, wie die Entwickler es tatsächlich implementierten. Auch die vom System erlaubten Aktionen waren nach dem Bestätigen der Konfiguration nicht die, welche ein Nutzer durchführen wollte. Im Kontext der HCI spricht man manchmal auch von einem Gulf of evaluation oder Gulf of execution, also einer Kluft zwischen den Aktionen, die ein Nutzer durchführen will, und denjenigen Aktionen, die das System erlaubt.

Vergleich der 5 Webschnittstellen

Weiter oben wurden die verschiedenen Typen von Webschnittstellen detailliert beschrieben. Zur Erinnerung: Diese Schnittstellen unterscheiden sich vor allem in der Platzierung und Funktionsweise von Navigationselementen und Suchfunktionen. Im Folgenden sollen nun die Ergebnisse der fünf Versuchsvarianten interpretiert und begründet werden.

Abbildung 4.25 vergleicht unterschiedliche Performance-Metriken. Werden diese zusammen betrachtet, so erlaubt dies unter der Berücksichtigung von Abbildung 4.24 eine Beurteilung, wie gut die unterschiedlichen Schnittstellen an die mentalen Modelle der Nutzer angepasst sind. Wie bereits erwähnt, scheint die vierte Schnittstelle die am wenigsten an die mentalen Modelle der Nutzer angepasste Nutzeroberfläche zu sein. Ein möglicher Grund dafür könnte die mangelnde externe Konsistenz (vgl. Kapitel 2, Abschnitt 2.3.3 Konsistenz) sein. Das bedeutet, dass die Schnittstelle die gewohnten Nutzungskonventionen der Nutzer nicht genügend berücksichtigt. Die Benutzer sind damit vermutlich nur beschränkt in der Lage, vorhandenes Wissen aus früheren Interaktionserfahrungen auf die Bedienung der Schnittstelle anzuwenden, was zu mehr Ineffizienz führte und die Erlernbarkeit erschwerte. Die subjektiven Beurteilungen der Nutzer unterstützen diese Behauptungen (vgl. Abbildung 4.27 und Abbildung 4.26).

Eine weitere Beobachtung, die sich aus Abbildung 4.25 ableiten lässt, ist, dass die Schnittstelle mit der Mega-Drop-Down-Menükomponente insgesamt am effizientesten genutzt wurde. Im Durchschnitt benötigten die Nutzer lediglich sieben Mausklicks zur erfolgreichen Bearbeitung der Aufgabe. Gleichzeitig zeigt sich jedoch, dass sie vergleichsweise mehr Zeit investierten, um den Suchprozess überhaupt zu starten (24.38 Sekunden im Schnitt). Dies deutet darauf hin, dass die Erlernbarkeit dieser Schnittstelle im Vergleich zu den anderen Schnittstellen möglicherweise geringer war.

Die hohe Effizienz ist vermutlich auf die Gestaltung des Mega-Drop-Down-Menüs zurückzuführen: Wird der Mauszeiger über einen Hauptmenüpunkt bewegt, werden alle verfügbaren Kategorien gleichzeitig angezeigt. Dadurch ist es möglich, mit nur einem Klick direkt zur gewünschten Produktkategorie zu navigieren. Al-

lerdings erfordert dieser Vorteil zunächst einen möglicherweise längeren visuellen Erkennungsprozess, um die richtige Kategorie zu identifizieren. Zudem ist anzunehmen, dass ein gewisser zeitlicher Aufwand notwendig war, um das Mega-Drop-Down-Menü überhaupt zu lokalisieren und seine Funktionsweise zu verstehen. Dies könnte erklären, weshalb der erste Klick im Mittel erst nach 24,38 Sekunden erfolgte.

Nutzerverhalten

Sowohl die Ergebnisse der Nutzerkonfigurationen (Abbildung 4.23) als auch die Analyse des Suchverhaltens (Abbildungen 4.28a und 4.28b) der Teilnehmer zeigen ein Muster: Je komplexer die Aufgabe (das heisst, je mehr Produktspezifikationen es gibt), desto seltener werden Suchfunktionen eingesetzt, um Produkte zu suchen. Im Gegensatz dazu steigt die Benutzung von textbasierten Menüs und komplexitätsreduzierenden Funktionen wie Filterfunktionen mit einer steigenden Anzahl an Produktspezifikationen an.

Dies ist möglicherweise eine Konsequenz aus dem Gesetz von Miller. Denn je mehr Produktspezifikationen verlangt werden, desto mehr Informationen müssen im Kurzzeitgedächtnis gespeichert und verarbeitet werden. Umso mehr kognitive Ressourcen werden also benötigt, um die Aufgabe zu bewältigen. Der Einsatz einer Suchfunktion (ohne Autovervollständigung) würde verlangen, die gesamten Informationen, die es braucht, um das Produkt zu finden, zeitgleich aus dem Kurzzeitgedächtnis abzurufen, um diese entsprechend in das Suchfeld eintippen zu können. Allenfalls würden nicht alle Spezifikationen abgerufen werden, sondern nur Suchbegriffe, welche eine erste Filterung in eine relevante Kategorie ermöglichen. Auch dies verlangt einen Erinnerungsprozess, um anschliessend solche Begriffe einzutippen. Hierfür würden ebenfalls mentale Ressourcen benötigt werden. Selbst wenn die Suchfunktion eine Autovervollständigung aufweisen würde, wäre zu Beginn ein Erinnerungsprozess nötig, um die korrekten (Teil)-Suchbegriffe einzutippen, um dann über einen Erkennungsprozess das gesuchte Produkt aus der Liste der vorgeschlagenen Produkte auszuwählen. Würde die Suchfunktion verwendet, so würde nach der Nutzung der Suchfunktion in aller Regel ein Erkennungs- oder Filterprozess erfolgen, um schliesslich das gewünschte Produkt, welches die Spezifikationen erfüllt, zu finden. Insgesamt benötigt diese Vorgehensweise also eine Kombination aus Erinnerungs- und Erkennungsprozessen und weist somit einen höheren Anteil an benötigten mentalen Ressourcen auf.

Der Einsatz eines textbasierten Menüs, bei dem die Produktkategorien über das Anklicken von Links aufgerufen werden, erfordert in erster Linie einen Erkennungsprozess – nämlich die Auswahl desjenigen Links, der zur gesuchten Produktkatego-

rie führt. Möglicherweise ist zu Beginn ein initialer Erinnerungsprozess notwendig, um sich die passende Kategorie ins Gedächtnis zu rufen. Allerdings kann der visuelle Hinweis durch die sichtbaren Links diesen Erinnerungsprozess erleichtern: Die angebotenen Optionen dienen als externe Gedächtnisstützen, wodurch Nutzende schneller erkennen können, welche Kategorie zielführend ist – insbesondere im Vergleich zu einer Situation ohne visuelle Unterstützung. Mit dem anschließenden Einsatz von Filterfunktionen, nachdem der Link betätigt worden ist, kann dann die Menge an Informationen weiter gezielt reduziert werden. Damit müssen weniger Informationen verarbeitet resp. gefiltert werden. Um das Produkt zu finden, sind also insgesamt weniger Erinnerungs- und Erkennungsprozesse notwendig, es werden also weniger mentale Ressourcen verlangt. Die Ergebnisse lassen also die Vermutung aufkommen, dass je mehr Informationen für das Bewältigen einer Aufgabe benötigt werden, umso mehr wird ein auf Erkennen und weniger auf Erinnern basiertes Vorgehen gewählt.

Ein letzter, bemerkenswerter Aspekt soll an dieser Stelle noch diskutiert werden. Ein genauerer Blick auf die Abbildungen 4.24 und 4.25 zeigt, dass die Teilnehmenden die Aufgabe im fünften Versuch – mit einer Schnittstelle, die ein Off-Canvas-Menü am rechten Bildschirmrand sowie eine Suchfunktion mit Autovervollständigung im unteren Bildschirmbereich enthält – mit einer vergleichbaren Performance bearbeiteten wie im ersten Versuch. In diesem war das Off-Canvas-Menü hingegen am linken Rand positioniert, während sich die Suchfunktion im oberen Bereich des Interfaces befand.

Besonders interessant ist in diesem Zusammenhang auch die Erkenntnis aus Abbildung 4.28a. Bereits zuvor wurde beobachtet, dass die Suchfunktion tendenziell häufiger genutzt wurde, je geringer die Aufgabenkomplexität ausfiel. Dieses Muster scheint jedoch im fünften Test ($T5$) nicht im gleichem Ausmass aufzutreten. Obwohl die Tests $T1$ und $T5$ identische Anforderungen hinsichtlich Anzahl und Art der Produktspezifikationen aufwiesen, wurde die Suchfunktion im fünften Versuch seltener eingesetzt.

Eine plausible Erklärung dafür könnte in der Platzierung der Suchleiste liegen. Es ist möglich anzunehmen, dass die Positionierung am unteren Bildschirmrand nicht dem mentalen Modell der Teilnehmenden entsprach. Infolgedessen wurde die Suchfunktion entweder nicht als solche wahrgenommen, nicht als hilfreich eingeschätzt oder nicht erkannt. Dies könnte dazu geführt haben, dass viele Nutzer nach alternativen Suchprozessen ausschau hielten. Hierbei spielt vermutlich das Off-Canvas-Menü eine zentrale Rolle (vgl. Abbildung 4.30). Es ist anzunehmen, dass das Menü-Icon durch seine visuelle Gestaltung und seinen starken Aufforderungscharakter (Affordance) viele Teilnehmende dazu motivierte, es zu verwenden und

ihre Produktsuche über die Menüstruktur durchzuführen. Trotz der untypischen Positionierung des Menüs am rechten Rand dürfte das Icon, ebenfalls aufgrund seines Aufforderungscharakters, schnell ins Blickfeld geraten sein. Dies könnte auch darauf hinweisen, dass durch den starken Aufforderungscharakter des Icons die Lücke zwischen dem mentalen Modell der Nutzer und der tatsächlichen Systemgestaltung teilweise überbrückt werden konnte. Diese Annahmen bieten auch eine Erklärung dafür, warum im fünften Versuch die Zeit bis zum ersten Mausklick auf das Menü kürzer ausfiel als in allen anderen Tests. Das Menü wurde offensichtlich schnell erkannt und seine Nutzung aufgrund der Gestaltung intuitiv verstanden.

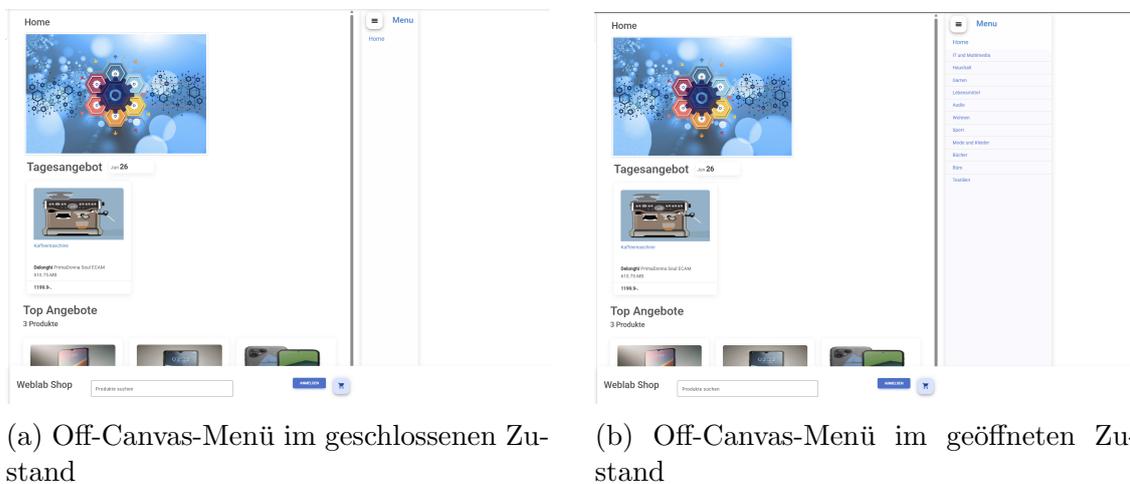
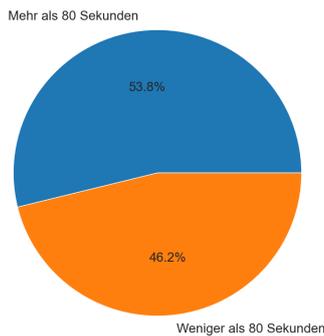


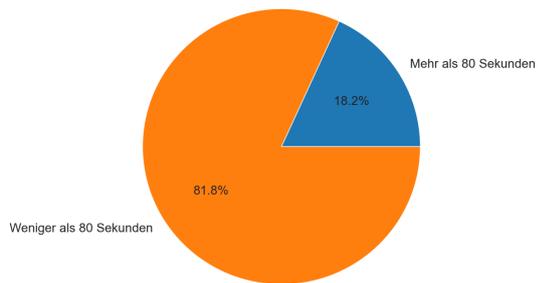
Abbildung 4.30: Off-Canvas-Menü im geschlossenen bzw. geöffneten Zustand

Obwohl Benutzer eines Systems meist ein mentales Modell besitzen, welches ihre bevorzugte Vorgehensweise massgebend festlegt, kann es manchmal sinnvoll sein, die Nutzer dennoch durch Schaltflächen mit starker Affordance oder modeless Feedback auf weitere, effizientere Vorgehensweisen hinzuweisen. Auf diese Weise können diese zumindest selbst entscheiden, welche Vorgehensstrategie sie wählen möchten. Der Grund hierfür zeigen Abbildung 4.23 und Abbildung 4.32. Die Abbildungen illustrieren, dass bei hoher Aufgabenkomplexität die Effizienz gesteigert werden kann, wenn Komponenten wie Filter oder textbasierte Menüs eingesetzt werden, die auf Erkennungsprozesse basieren – im Vergleich zu Funktionalitäten wie Suchfunktionen, die Erinnerungsleistungen erfordern. Der Grund hierfür ist vermutlich eine Konsequenz aus dem Gesetz von Miller und wurde bereits weiter oben erläutert. Aus diesem Grund ist es wichtig, Schnittstellen und ihre Funktionalitäten gut erkennbar zu gestalten, z.B. durch einen starken Angebotscharakter, damit die Nutzer alle möglichen Vorgehensweisen erkennen und die effizientesten wählen können.

Anzahl Nutzer mit Bearbeitungszeit weniger/mehr als 80 Sekunden mit Suchfunktion



Anzahl Nutzer mit Bearbeitungszeit weniger/mehr als 80 Sekunden ohne Suchfunktion



(a) Anteil Nutzer, welche mehr als 80 Sekunden benötigten, um Test 4 ($T4$) abzuschliessen

(b) Anteil Nutzer, welche mehr als 80 Sekunden benötigten, um Test 4 ($T4$) abzuschliessen

Abbildung 4.31: Anteil Nutzer, welche mehr oder weniger als 80 Sekunden benötigten, um Test 4 ($T4$) abzuschliessen

4.10 Restorff-Effekt

Auch der Restorff-Effekt und seine Wirkung auf die Usability und User Experience waren Gegenstand dieser Arbeit. Zu diesem Zweck wurden drei Versuchsvarianten entwickelt. In diesen Experimenten steht im Fokus, wie Nutzer ihre Aufmerksamkeit steuern und inwiefern eine Aufmerksamkeitslenkung gezielt beeinflusst oder herbeigerufen werden kann. Die Ergebnisse dieser Untersuchungen können allenfalls Erkenntnisse darüber liefern, wie sich die Aufmerksamkeit der Nutzer gezielt auf relevante Inhalte lenken lässt und in welchem Ausmass solche Massnahmen unterstützend wirken, bevor sie als störend oder ablenkend empfunden werden und die Usability nicht mehr verbessern.

4.10.1 Testaufbau

Um die Auswirkungen des Restorff-Effekts auf die Usability und die Aufmerksamkeitslenkung zu untersuchen, wurden drei Versuchsvarianten entwickelt. Alle drei Varianten basieren auf einer grafischen Oberfläche einer E-Mail-Anwendung, die eine Liste von E-Mails im Posteingang verwaltet. Jede E-Mail verfügt am rechten Rand über eine Gruppe von drei Schaltflächen. Die erste Schaltfläche (von links) dient dem Weiterleiten der E-Mail, über die mittlere kann die E-Mail gelöscht werden, und die rechte Schaltfläche ermöglicht das Bearbeiten der E-Mail.

Der Ablauf ist in allen drei Testvarianten identisch: Sobald eine Versuchsperson den Test startet, wird in regelmässigen Abständen eine neue E-Mail zufällig am Anfang oder am Ende des Posteingangs eingefügt. Die Aufgabe der Teilnehmenden besteht darin, diese E-Mail unmittelbar nach dem Erscheinen entweder zu löschen

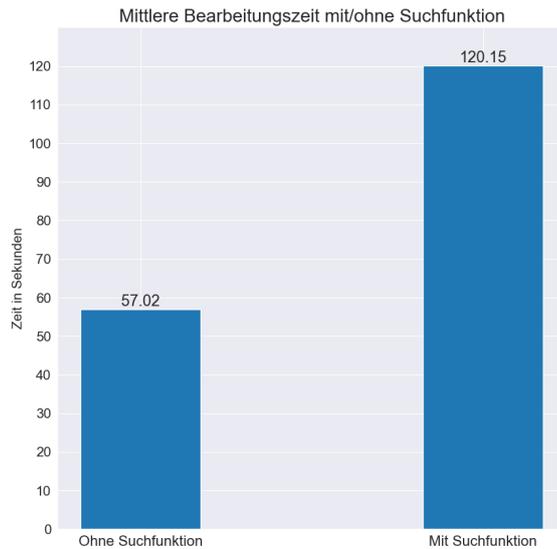


Abbildung 4.32: Mittlere Bearbeitungszeit des vierten Tests (T_4) mit und ohne Suchfunktion

oder weiterzuleiten. Welche Aktion jeweils erforderlich ist, wird vor dem Hinzufügen der E-Mail über ein entsprechendes Hinweissfeld rechts neben der E-Mail-Liste angezeigt. Nachdem eine Aktion ausgeführt wurde, wird die Instruktion für die nächste E-Mail aktualisiert. Insgesamt werden auf diese Weise sechs E-Mails in die Liste eingefügt.

Die Gestaltung der Schaltflächen zum Bearbeiten, Löschen und Weiterleiten unterscheidet sich zwischen den drei Varianten: In der ersten Variante unterscheiden sich die drei Aktionsflächen farblich nicht. Es werden keine Kontraste eingesetzt, um das Weiterleiten, Löschen oder Bearbeiten farblich zu trennen. Bewusst wird hier also keine Aufmerksamkeit auf die Zielflächen gelenkt, der Restorff-Effekt wird also nicht oder kaum berücksichtigt. In der zweiten Variante des Versuches werden die Aktionsflächen ebenfalls nicht farblich hervorgehoben oder getrennt. Neu wird der Betreff der hinzugefügten E-Mail aber farblich hervorgehoben, so dass er sich von allen anderen Textinhalten farblich unterscheidet. In der dritten Versuchsvariante wird die Schaltfläche der Löschfunktion farblich von den anderen beiden Funktionen (Weiterleiten und Bearbeiten) unterschieden. Die Lösch-Schaltfläche ist nun blau gefärbt und das Icon in der Mitte der Schaltfläche sticht mit einer weissen Farbe hervor. Die anderen Schaltflächen für das Bearbeiten und Löschen bleiben in derselben Farbe wie in der ersten Variante. Die farbliche Hervorhebung des E-Mail-Betreffs bleibt ebenfalls erhalten. Die dritte Testvariante weist somit die stärksten visuellen Kontraste auf. Die grafischen Darstellungen der drei Versuchsvarianten sind im Anhang D, Abschnitt D.9, Restorff-Effekt, abgebildet.

4.10.2 Usability und User Experience-Metriken

Performance-Metriken eignen sich für die Untersuchung der Auswirkungen des Restorff-Effektes auf die Usability und Performance der Nutzer. Als Performance-Metrik wird die Aufgabenerfolgsquote (Task-Success) herangezogen. Dies hilft zu analysieren, ob die jeweiligen Versuche und einzelnen Aufgaben innerhalb der Versuche vollständig und korrekt gelöst wurden.

4.10.3 Ergebnisse

Abbildung 4.33 zeigt die Erfolgsquoten nach Versuchsvariante. Die Grafik zeigt, dass der erste Versuch die geringste Erfolgsquote (etwa 27%) aufweist und der zweite und dritte Versuch eine beinahe gleiche Erfolgsquote besitzen, ca. 50%.

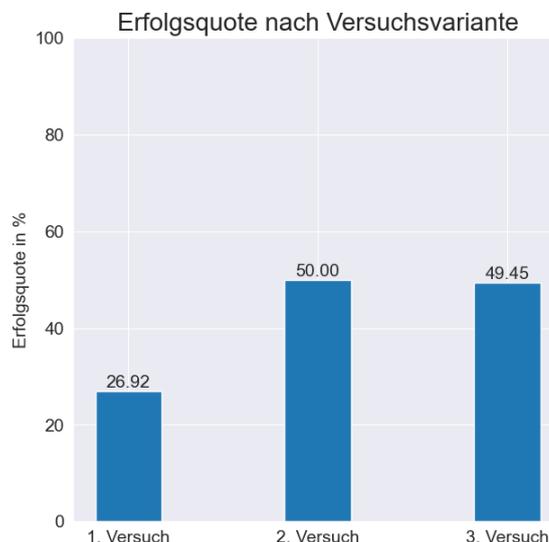
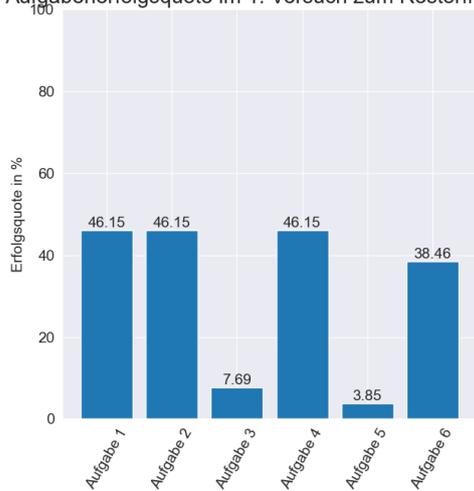


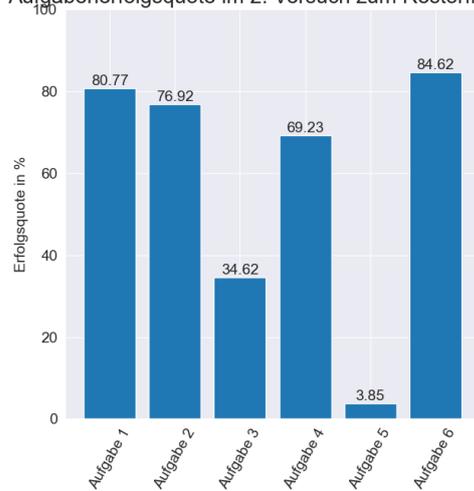
Abbildung 4.33: Erfolgsquote nach Versuchsvariante

Abbildung 4.34 und Abbildung 4.35 zeigen die Erfolgsquoten je Aufgabe in jedem Versuch an (insgesamt 6 Aufgaben). Die geringsten Erfolgsquoten weisen die einzelnen Aufgaben im ersten Versuch auf. Zur Erinnerung: Im ersten Versuch wurden keine farblichen Hervorhebungen oder Kontraste eingesetzt, um die Aufmerksamkeit der Nutzer gezielt zu lenken. Im zweiten Versuch wurde der Betreff der E-Mail farblich hervorgehoben und im dritten Versuch unterschied sich der Lösch-Button farblich von den anderen zwei Buttons. Im zweiten und dritten Versuch fallen die Erfolgsquoten bereits höher aus. In diesen beiden Versuchstests fallen die Erfolgsquoten ähnlich aus. Die geringsten Erfolgsquoten besitzen die dritte und fünfte Aufgabe, wobei die Quoten für die restlichen Aufgaben alle höher als 50% ausfallen.

Aufgabenerfolgsquote im 1. Versuch zum Restorff-Effekt



Aufgabenerfolgsquote im 2. Versuch zum Restorff-Effekt



(a) Erfolgsquote der einzelnen Aufgaben im ersten Versuch zum Effekt von Restorff

(b) Erfolgsquote der einzelnen Aufgaben im ersten Versuch zum Effekt von Restorff

Abbildung 4.34: Erfolgsquote der einzelnen Aufgaben im ersten und zweiten Versuch zum Effekt von Restorff

Aufgabenerfolgsquote im 3. Versuch zum Restorff-Effekt

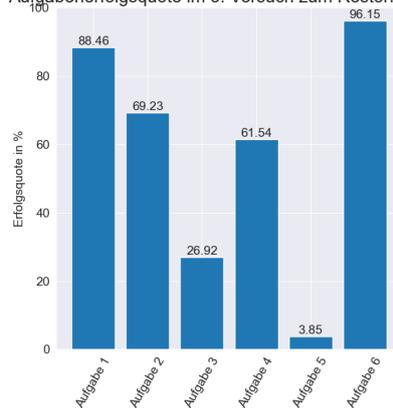


Abbildung 4.35: Erfolgsquote je Aufgabe im dritten Versuch zum Effekt von Restorff

4.10.4 Interpretation der Ergebnisse

Bei der Interpretation der Ergebnisse soll auf übermäßige Spekulation verzichtet werden. Es erscheint plausibel, dass die schrittweise intensivere Berücksichtigung des Restorff-Effekts – realisiert durch farbliche Hervorhebungen und Kontraste zwischen den Schaltelementen – einen positiven Einfluss auf die Performance hatte. Diese Massnahmen lenkten die visuelle Aufmerksamkeit auf zentrale Bedienelemente, wodurch Fehlclicks reduziert, die Effizienz erhöht und die Erfolgsquote insgesamt verbessert werden konnte. Der Einsatz des Restorff-Effekts trug somit zur Steigerung der Usability bei.

Der stärkste Einsatz des Restorff-Effekts erfolgte im dritten Versuchsdurchlauf.

Allerdings zeigte sich, dass dieser im Vergleich zum zweiten Versuch weder zu einer signifikant höheren noch zu einer niedrigeren Erfolgsquote führte (vgl. Abbildung 4.33). Die zusätzliche visuelle Hervorhebung wurde demnach vermutlich nicht als ablenkend wahrgenommen, konnte die Nutzer in der Zielfindung jedoch auch nicht weitergehend unterstützen. Abbildung 4.34 und Abbildung 4.35 zeigen, dass die dritte und fünfte Aufgabe in allen drei Versuchsbedingungen die geringsten Erfolgsquoten aufweisen. Eine plausible Erklärung dafür liegt vermutlich an der Aufgabe selbst: Während in den Aufgaben eins, zwei, vier und sechs jeweils die neu hinzugefügte E-Mail gelöscht werden musste, bestand die Handlung in Aufgabe drei und fünf darin, die E-Mail weiterzuleiten – es musste also eine andere Schaltfläche als die Löschfunktion verwendet werden. Obwohl die Instruktionen entsprechend angepasst wurden, lag der Aufmerksamkeitsfokus vieler Nutzer während der Tests primär auf der Lokalisierung der neu hinzugefügten E-Mail und dem reflexartigen Ausführen des Löschvorgangs. Das aufmerksame Lesen der veränderten Instruktion wurde dadurch offenbar vernachlässigt.

Kapitel 5

Schlussfolgerungen und künftige Arbeiten

In diesem Kapitel werden die wichtigsten Erkenntnisse der Arbeit zusammengefasst sowie mögliche künftige Arbeiten vorgestellt.

5.1 Schlussfolgerungen

Das Ziel der Arbeit war es, durch einen Usability- und User Experience-Test mit 37 Teilnehmenden empirisch zu untersuchen, wie sich verschiedene UX-Gesetze auf die Usability und insgesamt auf die User Experience von Webschnittstellen auswirken. Die wichtigsten Erkenntnisse dieser Untersuchung werden im folgenden Abschnitt zusammengefasst.

Der im Rahmen dieser Arbeit durchgeführte Usability- und User Experience-Test kombiniert Elemente eines klassischen, moderierten sowie unmoderierten Tests. Während des Tests interagierten die Teilnehmenden selbständig mit einer voll funktionsfähigen Webseite, auf der sie alle vorgesehenen Experimente durchführten und dabei eine individuelle User Experience erzeugten. Im Zentrum stand unter anderem das Gesetz von Miller, dem zufolge jeder Interaktionsprozess mentale Ressourcen beansprucht. Nutzer müssen dabei Informationen entweder aus dem Kurzzeit- oder Langzeitgedächtnis abrufen, speichern oder verarbeiten. Je höher die kognitive Belastung durch eine Interaktion ist, desto negativer kann sich dies auf die Effizienz auswirken. Ein entscheidender Faktor ist, ob eine Aufgabe überwiegend Erinnerungs- oder Erkennungsprozesse erfordert.

Die Untersuchungen zu Millers Gesetz zeigten, dass dessen Berücksichtigung einen Einfluss auf die Nutzerperformance haben kann. Interaktionen, die primär auf Erinnerungsleistungen beruhen, wurden von den Teilnehmenden zwar nicht als

belastender wahrgenommen, benötigten jedoch im Schnitt mehr Zeit als Aufgaben, die auf Erkennen basierten. Dies liegt vermutlich daran, dass Erinnerungsprozesse oft mit zusätzlichen kognitiven Anforderungen verbunden sind, insbesondere wenn sie mit Erkennungsaufgaben kombiniert werden. Die Resultate legten auch dar, dass initiale Lernprozesse die Interaktionszeit zu Beginn erhöhen können. Wurde die Schnittstelle anschliessend erlernt, konnte die Effizienz gesteigert werden. Diese Tatsache konnte einerseits während der beiden Untersuchungen zum Prinzip der Kontrolle und Steuerbarkeit untersucht werden und andererseits war der erste Erkennungsversuch in der Experimentreihe „Erinnern vs. Erkennen“ für viele Teilnehmende der allererste Versuch, was zu einer erhöhten Verzögerung vor dem ersten Klick, der den Handlungszyklus einleitete, führte. Auch die Untersuchungen zum Gesetz von Jakob und zu mentalen Modellen zeigten, dass Benutzeroberflächen, die weniger auf verbreitete Designkonventionen zurückgreifen oder komplexere Elemente wie etwa ein Mega-Drop-Down-Menü enthalten, mehr Zeit zum Erlernen erfordern als solche, die auf gängige Designnormen setzen.

Gleichzeitig wurde aber auch deutlich, dass gezielt eingesetzte Funktionen, die auf Erinnern basieren, wie zum Beispiel Suchfunktionen, bei gleichbleibend niedriger Komplexität zu effizienten Nutzungsergebnissen führen können. Dies deutet darauf hin, dass die Gestaltung nicht zwingend ausschliesslich auf Erkennung setzen muss, sondern dass auch Erinnerungsprozesse effektiv sein können – sofern sie sinnvoll eingebettet und gut unterstützt werden.

Die fehlende Berücksichtigung von Millers Gesetz bei der Darstellung von Textinformationen, etwa bei Feedback in Form von Dialogen, hatte spürbare Auswirkungen auf die Ergebnisse zur Untersuchung des Paradoxons des aktiven Nutzers. Der Dialog, der den Nutzenden zu Beginn der User Experience angezeigt wurde, enthielt gemäss den Ergebnissen zu viel Text auf einmal. Dadurch konnten sich die Nutzer den Inhalt nicht genügend merken, das Kurzzeitgedächtnis war also kurzzeitig überfordert, und sie folgten den Vorschlägen des Dialogs nicht. Wenn Nutzende über Funktionen wie Einstellungen oder Hilfestellungen informiert werden sollen, ist es deshalb wichtig, dass Feedback klar, prägnant und leicht verständlich formuliert ist. Ein gutes Beispiel dafür war der Hinweis-Dialog, der nach Abschluss der zweiten Kategorie eingeblendet wurde. Er beschränkte sich auf einen einzelnen Hinweis zu den Einstellungen und war einfach formuliert. Dieses Feedback blieb bei den Nutzenden deutlich besser im Gedächtnis und führte dazu, dass einige die Einstellungen aufriefen und Änderungen vornahmen.

Die Untersuchungen zum Gesetz von Tesler zeigten, dass eine gezielte Benutzerführung sowie die Verlagerung von Entscheidungsverantwortung auf das System die Interaktion vereinfachen, die Effizienz der Nutzer steigern und deren kognitive

Belastung verringern können. In dieser Arbeit wurde eine solche Benutzerführung durch den Einsatz der sogenannten progressiven Offenlegung bei der Auswahl der Experimente umgesetzt. Die Ergebnisse verdeutlichten zudem, dass bei einigen Teilnehmenden das Fehlen einer strukturierten Benutzerführung sowie die damit verbundene erhöhte Komplexität bei Entscheidungs- und Auswahlprozessen aus Effizienzperspektive zu Usability-Problemen führen kann. Daher ist es empfehlenswert, Entscheidungsprozesse auf das System zu übertragen – vorausgesetzt, diese sind nicht individuell notwendig oder kritisch. Auf diese Weise lässt sich die kognitive Belastung reduzieren, die Effizienz steigern und eine klarere Nutzerführung unterstützen.

Sehr ausführlich wurde das Gesetz von Jakob und mentale Modelle von Nutzern unterschiedlicher Webschnittstellen untersucht. Während der Nutzerkonfigurations-tests konnte festgestellt werden, dass Nutzer vermutlich bereits über ein mentales Modell eines Webshops (Nutzungskontext) verfügen. Entsprechend diesem mentalen Modell wurden die Webshops konfiguriert, ohne dabei gezielt darauf zu achten, welches Produkt tatsächlich gesucht werden musste. Allerdings zeigten sich Diskrepanzen zwischen den erwarteten Funktionen und der tatsächlichen Nutzung, abhängig von der Aufgabe (Problemkontext). Für die Gestaltung von Produkten sollte in Betracht gezogen werden, Systeme so einfach wie möglich, aber so funktional wie nötig zu gestalten. Die Herausforderung ist dabei, die Komplexität gemäss Teslers Gesetz zu minimieren, ohne von den Erwartungen (mentalen Modellen) der Zielgruppe abzuweichen. Obwohl Nutzer über ein mentales Modell verfügen, das ihre Vorgehensweise in gewissem Masse vorgibt, ist es dennoch wichtig, Systeme so zu gestalten, dass sie über eine hohe Selbstbeschreibungsfähigkeit verfügen und ihre verfügbaren Funktionen klar kommunizieren. So werden Nutzer auf möglicherweise effizientere Vorgehensweisen hingewiesen. Ebenso ist es in der Regel erforderlich, mehrere Iterationen von Usability-Tests durchzuführen, um das Design gezielt an die mentalen Modelle der Zielgruppe anzupassen.

Als letztes wurde der Restorff-Effekt und dessen Auswirkungen auf die Usability untersucht. Die Ergebnisse zeigten deutlich, dass bereits eine minimale Berücksichtigung dieses Effektes durch den Einsatz von farblichen Unterschieden und Kontrasten die Usability durch eine gezielte Lenkung der Aufmerksamkeit der Nutzer auf Zielelemente verbesserte.

5.2 Künftige Arbeiten

Abschliessend werden nun noch mögliche künftige Arbeiten vorgestellt. Das Alter der Probanden wurde in den experimentellen Untersuchungen dieser Arbeit nicht

berücksichtigt. Künftige Tests könnten entweder mit derselben Teilnehmergruppe oder mit einer neuen Gruppe wiederholt werden, wobei die Teilnehmenden in Altersgruppen unterteilt würden. Dies würde ermöglichen, zu untersuchen, ob und inwiefern das Alter einen Einfluss auf die Interaktionseffizienz hat und inwiefern sich die mentalen Modelle älterer Nutzerinnen und Nutzer von denen jüngerer unterscheiden.

Ein weiterer möglicher Untersuchungsaspekt betrifft die mentalen Modelle der Nutzerinnen und Nutzer. In den Ergebnissen sowie in der Schlussfolgerung wurde aufgezeigt, dass es gemäss Teslers Gesetz herausfordernd ist, die Komplexität einer Schnittstelle möglichst stark zu reduzieren, ohne dabei das mentale Modell der Nutzenden zu vernachlässigen. Eine weiterführende Arbeit könnte experimentell untersuchen, in welchem Ausmass Funktionalitäten reduziert oder umgestaltet werden können, ohne die Übereinstimmung mit dem mentalen Modell oder die Nutzerzufriedenheit negativ zu beeinflussen.

Die Ergebnisse brachten hervor, dass zu viele Informationen in Form von Text in Feedbacks, wie zum Beispiel in Dialogen, das Kurzzeitgedächtnis der Nutzer potenziell überbeanspruchen und die Informationen nicht haften bleiben. Eine Fortführung der Arbeit könnte untersuchen, welche Textmengen und -formen für Feedback oder Systemerklärungen optimal sind, indem unterschiedliche Varianten von Feedbacks eingesetzt werden.

Weitere experimentelle Untersuchungen könnten sich damit befassen, wie stark das Abweichen von Konventionen – etwa im Bereich der Navigation – die User Experience beeinflusst, insbesondere bei Nutzergruppen mit geringer digitaler Erfahrung oder höherem Alter. In diesem Zusammenhang könnten weitere Tests durchgeführt werden, um zu untersuchen, mit welchen Gestaltungsprinzipien, wie zum Beispiel Feedbackmechanismen oder ein Aufforderungscharakter, mögliche Lücken zwischen den mentalen Modellen von Entwicklern und jenen der Nutzenden überbrückt werden können.

Anhang A

Illustrationen zu UX-Gesetzen und Gestaltungsprinzipien

A.1 Gesetz von Miller

Die Anwendung des Gesetzes von Miller lässt ein Text beispielsweise hierarchisch, strukturiert und formatiert erscheinen, was beim Lesen zu einer geringeren kognitiven Belastung führt.

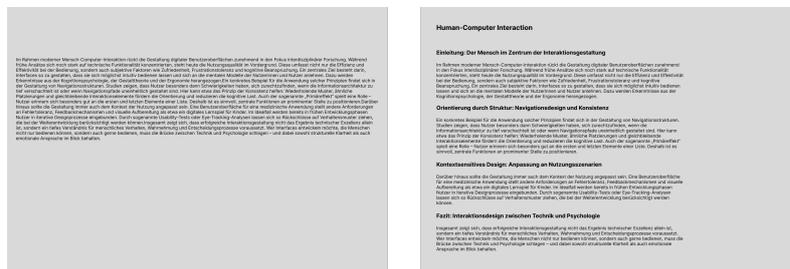


Abbildung A.1: Anwendung von Millers Gesetz auf einen Textabschnitt

A.2 Gesetz von Jakob

Das schrittweise Einführen neuer Versionen sowie den Benutzern während einer gewissen Zeitspanne die Möglichkeit geben, zur alten Version zurückzugehen, sind direkte Anwendungen des Gesetzes von Jakob.

A.3 Mentale Modelle

Abbildung A.3 ist ein Design einer Webschnittstelle, welches das mentale Modell von Nutzern einer Webanwendung durch die Berücksichtigung bekannter Designkonven-

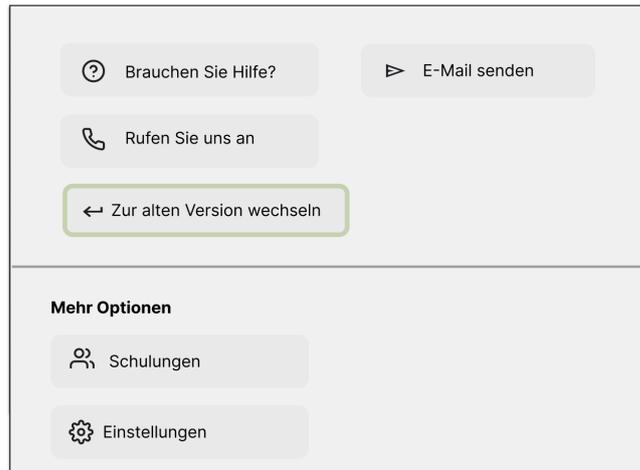


Abbildung A.2: Die Möglichkeit, zur vorherigen Version zurückzukehren, stellt eine Anwendung des Gesetzes von Jakob dar

tionen möglicherweise berücksichtigt.

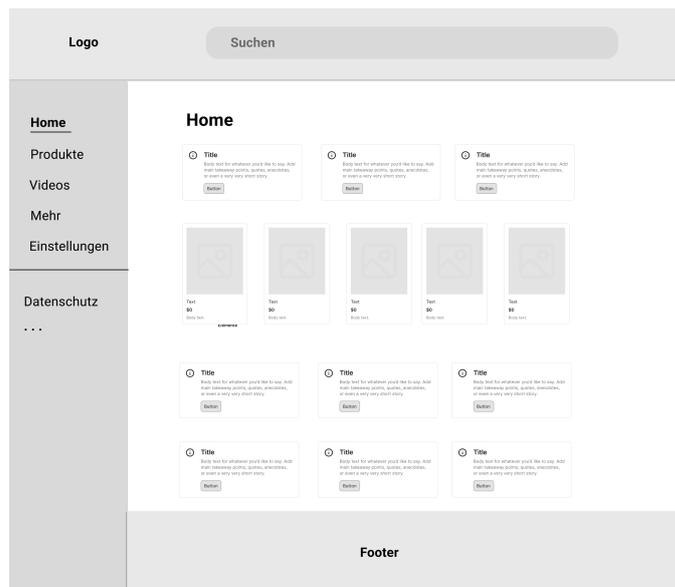


Abbildung A.3: Ein Design einer Webschnittstelle mit einer horizontalen Navigationsleiste am oberen Bildschirmbereich und einem Navigationsmenü am linken Bildschirmbereich. Das Design berücksichtigt Designkonventionen und ist damit möglicherweise an das mentale Modell von Webseitenutzer angepasst

A.4 Restorff-Effekt

Der Restorff-Effekt hilft, die Aufmerksamkeit auf zielführende Elemente zu lenken und Fehlklicks zu vermeiden.

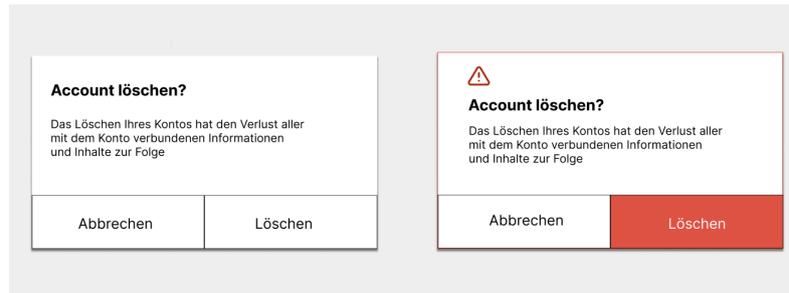


Abbildung A.4: Zwei Beispiele für Dialoge: Der linke Dialog berücksichtigt den Restorff-Effekt kaum, während der rechte diesen durch gezielte farbliche Hervorhebungen umsetzt.

A.5 Steuerbarkeit und Kontrollierbar

Abbildung A.5: Ein Formular, welches erlaubt selbst zu entscheiden, in welcher Reihenfolge die Eingaben abzuschliessen. Über die Links am oberen Bereich können Kategorien aufgerufen werden und Eingaben angepasst werden

A.6 Flexibilität und mentale Modelle

A.7 Angebotscharakter

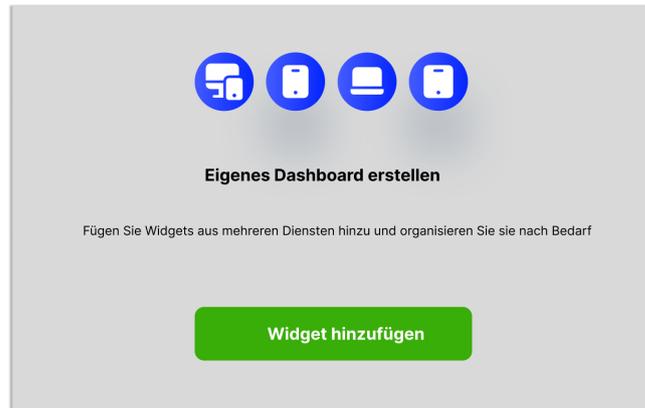


Abbildung A.6: Können Nutzer selbst Inhalte erstellen, so wird ihnen erlaubt, das System gemäss ihrem mentalen Modell aufzubauen

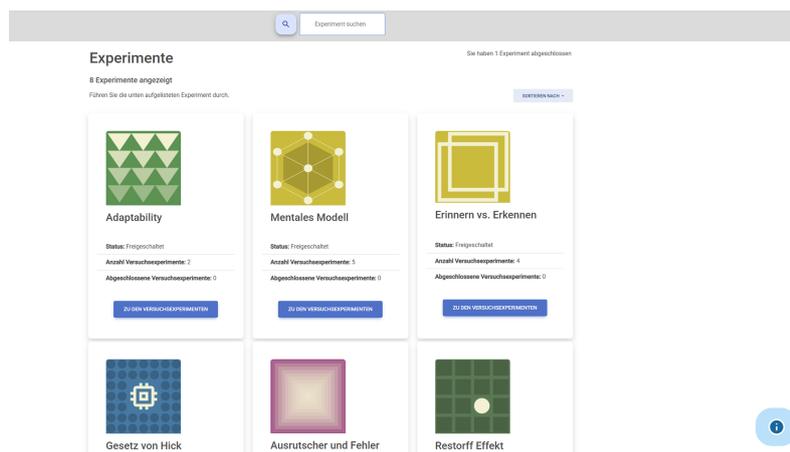


Abbildung A.7: Die Gestaltung des runden blauen Buttons weist einen starken Angebotscharakter auf. Durch die Form und Struktur weist der Button darauf hin, dass er betätigt werden kann oder soll

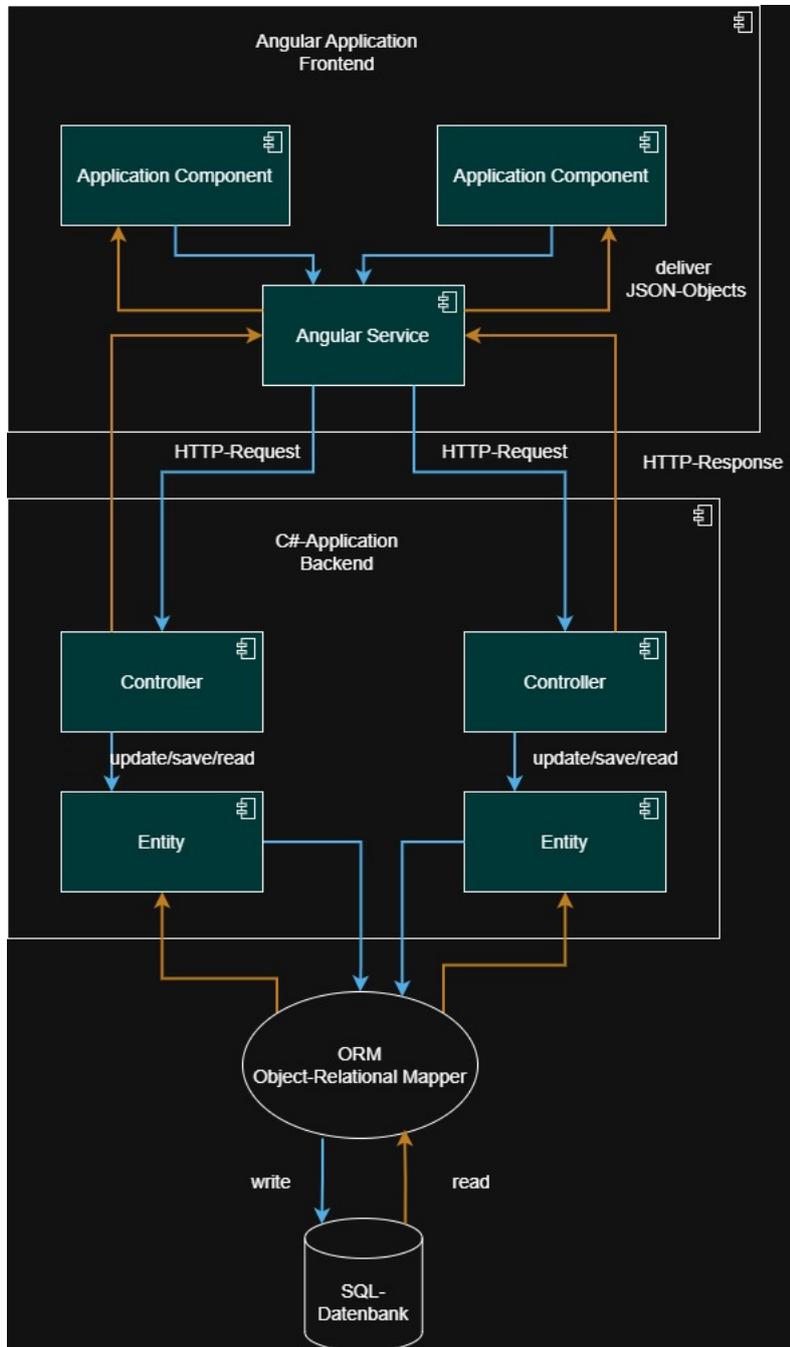


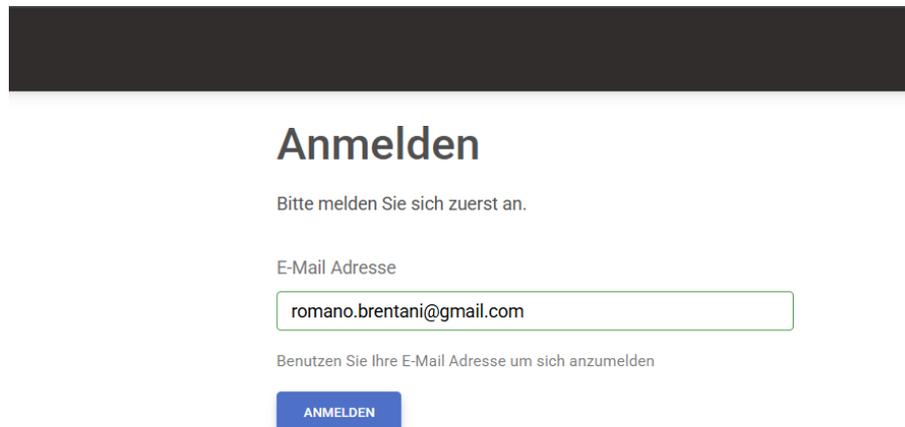
Abbildung B.1: Schichtendiagramm der Webseite

Anhang C

Webseite - Screenshots

C.1 Anmeldeformular

Das Anmeldeformular, über das sich Nutzer mit ihrer E-Mail-Adresse zu Beginn der User Experience anmelden.



The screenshot shows a login form with the following elements:

- Title:** Anmelden
- Instruction:** Bitte melden Sie sich zuerst an.
- Label:** E-Mail Adresse
- Input Field:** Contains the email address romano.brentani@gmail.com.
- Instruction:** Benutzen Sie Ihre E-Mail Adresse um sich anzumelden
- Button:** ANMELDEN

Abbildung C.1: Anmeldeformular

C.2 Umfrageformular

Nach dem Abschluss einiger Versuchstests sind die Teilnehmer gefragt, das Umfrageformular auszufüllen.

The screenshot shows a survey form titled "Schlussfeedback" within a "Weblab" application. The form contains five numbered questions, each with radio button options for responses. The questions are:

1. Erforderte das Durchführen des Experimentes viel Mühe, Anstrengung, Aufmerksamkeit oder Konzentration?
Options: 1 sehr gering, 2 gering, 3 mittel, 4 hoch, 5 sehr hoch
2. Funktionierte die grafische Oberfläche, welche Sie oben benutzten, genauso Ihrem mentalen Modell, d.h., funktioniert die Oberfläche so, wie Sie das erwartete Verhalten ist, wie Sie dies von anderen Modellen her kennen?
Options: 1 sehr gering, 2 gering, 3 mittel, 4 hoch, 5 sehr hoch
3. War die Funktionsweise der grafischen Oberfläche klar und verständlich?
Options: 1 kaum verständlich, 2 gering verständlich, 3 angemessen verständlich, 4 gut verständlich, 5 sehr verständlich
4. War die Funktionsweise der grafischen Oberfläche einfach erlernbar, strukturiert und erkennbar?
Options: 1 sehr schwierig, 2 schwierig, 3 angemessen einfach, 4 einfach, 5 sehr einfach
5. Waren die Anweisungen, Schritte und die Funktionsweise der Schnittstelle konsistent, d.h. befolgte die Schnittstelle Muster und wussten diese Muster über die gesamte Interaktion hinweg beizubehalten?
Options: 1 kaum konsistent, 2 wenig konsistent, 3 angemessen konsistent, 4 gut konsistent, 5 sehr konsistent

Below the questions is a text input field for additional feedback: "Gibt es Funktionslücken oder Schwächen, welche die Schnittstelle nicht zur Verfügung stellt, welche Sie aber gerne in der Schnittstelle gehabt hätten?". A green "Abschicken" button is at the bottom.

Abbildung C.2: Umfrageformular für die Nutzerbewertungen

C.3 Navigationsstrukturen

C.3.1 Linke Seitennavigation (Off-Canvas-Menü)

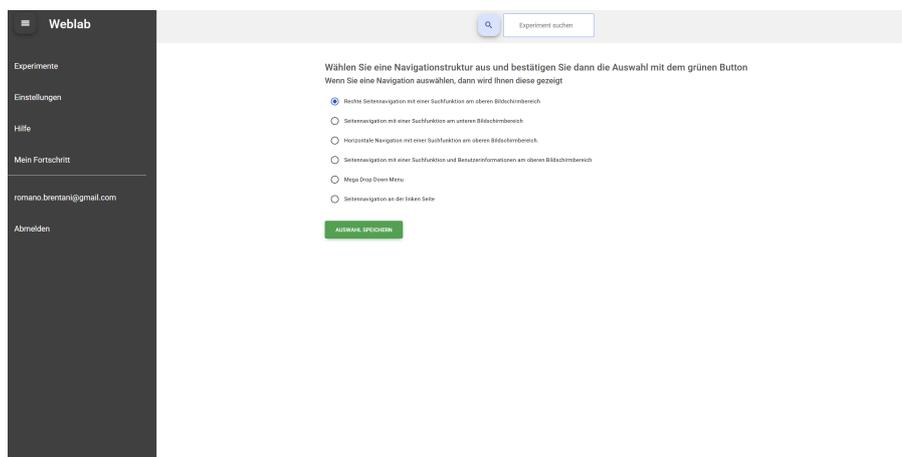


Abbildung C.3: Off-Canvas-Navigationsmenü am linken Bildschirmbereich

C.3.2 Horizontale Navigation

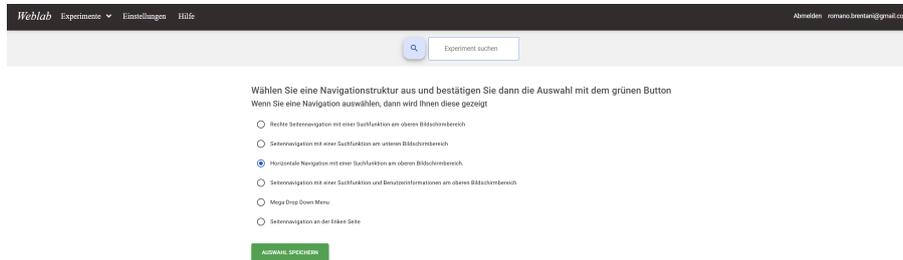


Abbildung C.4: Schichtendiagramm der Webseite

C.3.3 Meg-Drop-Down Navigation

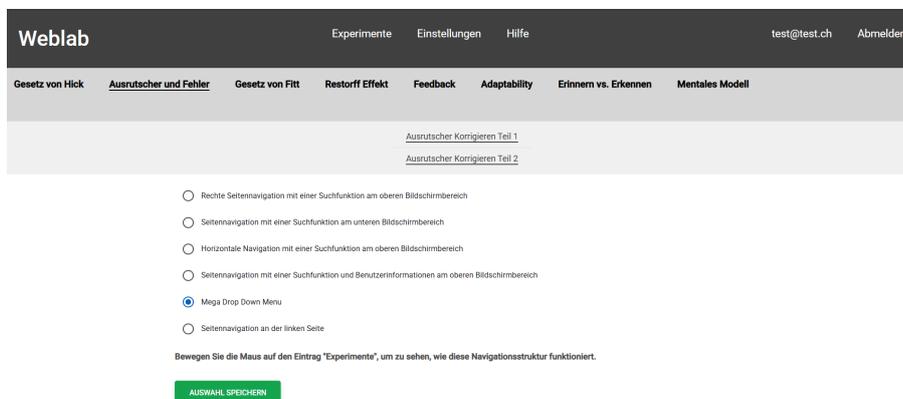


Abbildung C.5: Mega-Drop-Down Navigation

C.4 Hauptseite der Webseite

Die Hauptseite (Main-Page) zeigt alle Experimentkategorien an.

C.4.1 Experimente

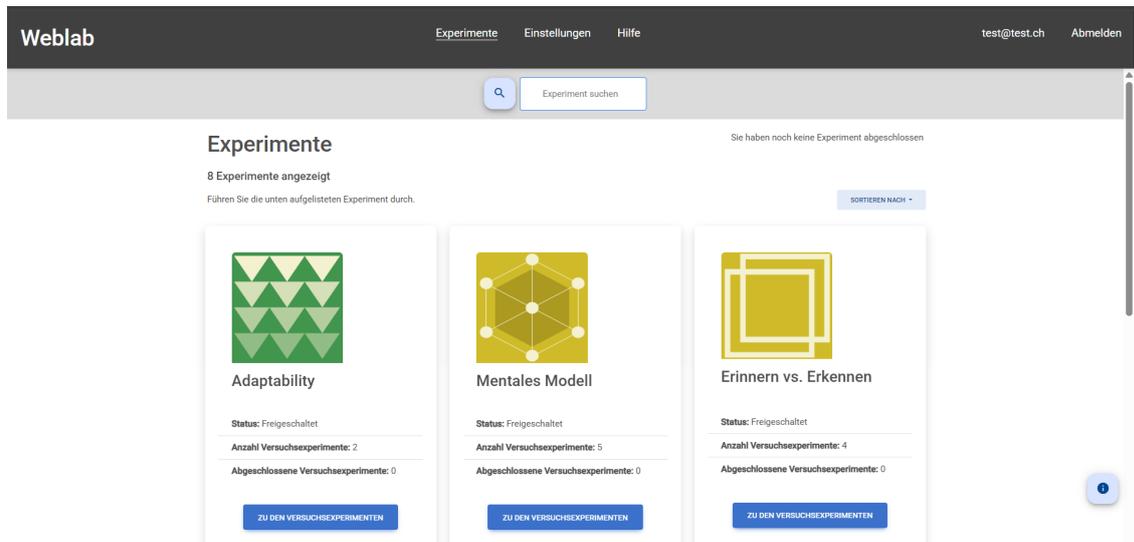


Abbildung C.6: Liste der Kategorien

C.4.2 Fortschrittsanzeige

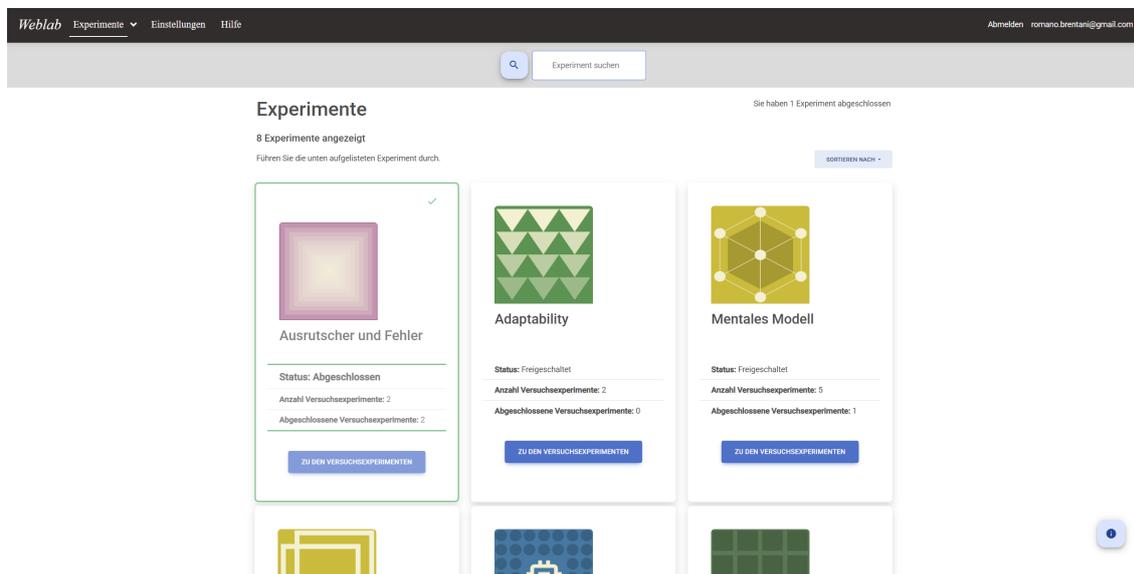


Abbildung C.7: Eine abgeschlossene Kategorie wird grün umrandet und die farblichen Effekte etwas reduziert um zu signalisieren, dass diese abgeschlossen wurde

C.5 Einstellungen

Die Einstellungen ermöglichen das Aktivieren der progressiven Offenlegung sowie des automatischen Wählens eines Versuchstests, sobald eine Kategorie ausgewählt wurde.



Abbildung C.8: Link, um zu den Einstellungen zu navigieren

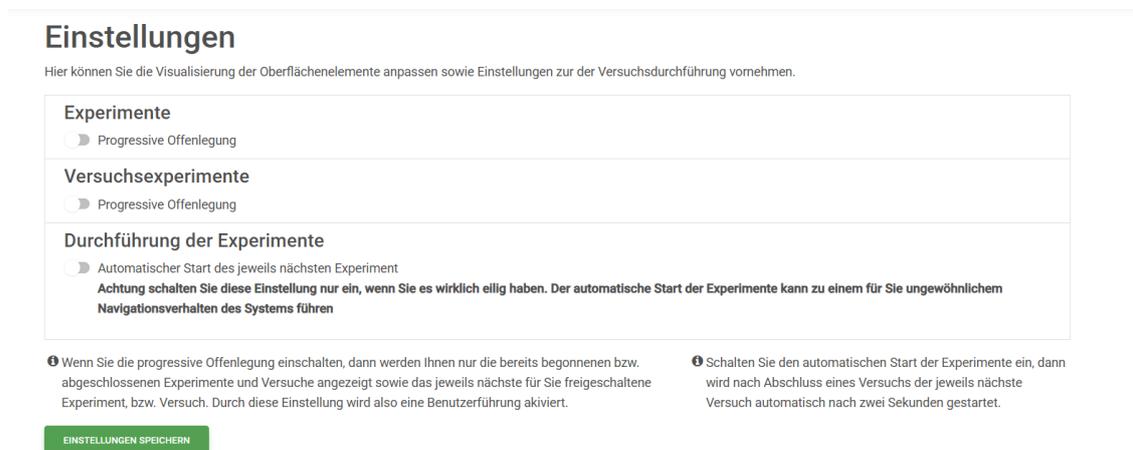


Abbildung C.9: Einstellungsoptionen der Webseite

C.6 Versuchsliste

Die Übersichtsliste innerhalb einer Kategorie zeigt alle verfügbaren Versuche einer Experimentkategorie an. Über das Betätigen des blauen Buttons wird ein Versuch ausgewählt.

The screenshot shows a web interface for an experiment category. At the top, there is a search bar with a magnifying glass icon and the text 'Versuche suchen'. Below this, the category title 'Erinnern vs. Erkennen' is displayed, followed by the instruction 'Bitte führen Sie die unten aufgelisteten Versuche durch.' To the right of the title is a yellow square graphic with a white border. Below the title, it states 'Total 4 Versuche (4 Versuche angezeigt)' and '0 Versuche von 4 Versuchen abgeschlossen.' The main content area contains two experiment cards. Each card has a title, a subtitle, a status 'Dieser Versuch ist freigeschaltet', a description, a status 'Status: Freigeschaltet', and an estimated duration. At the bottom of each card is a blue button labeled 'ZU DEN VERSUCHSDetails'.

Erinnern vs. Erkennen

Bitte führen Sie die unten aufgelisteten Versuche durch.

Total 4 Versuche (4 Versuche angezeigt)

0 Versuche von 4 Versuchen abgeschlossen.

1. Versuch
Erinnern vs. Erkennen Teil 1
Dieser Versuch ist freigeschaltet

In diesem Versuch geht es um das Erkennen von Funktionalitäten und Zuständen einer grafischen Oberfläche einer Web-Schnittstelle. Ihre Aufgabe ist es, ein Notebook Ihrer Lieblingsmarke zu finden und in den Warenkorb zu legen.

Status: Freigeschaltet

geschätzte Durchführungszeit: 2 Minuten

ZU DEN VERSUCHSDetails

2. Versuch
Erinnern vs. Erkennen Teil 2
Dieser Versuch ist freigeschaltet

In diesem Versuch geht es ebenfalls um das Erkennen oder Erinnern von Systemzuständen und Funktionalitäten einer grafischen Oberfläche. Ihre Aufgabe ist es, eine Gemüsesorte welche Sie mögen zu finden und in den Warenkorb zu legen. Nachdem Sie eine Sorte ausgewählt haben, müssen Sie zur Kasse gehen.

Status: Freigeschaltet

geschätzte Durchführungszeit: 1 Minuten

ZU DEN VERSUCHSDetails

Abbildung C.10: Übersichtsliste der Versuche innerhalb einer Experimentkategorie

C.7 Versuchsdetails

Eine detaillierte Beschreibung des Versuchs wird vor dem Start präsentiert.

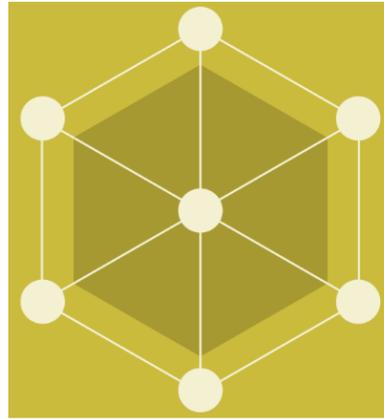
Erlernbarkeit und Mentale Modelle Teil 5

Versuchsbeschreibung

Erlernbarkeit und Mentale Modelle Teil 5

Suchen und Finden Sie das spezifizierte Produkt. Die Produktspezifikationen werden Ihnen in der Versuchsanleitung angegeben, sobald Sie den Versuch gestartet haben.

In diesem Versuch geht es um die Erlernbarkeit einer grafischen Oberfläche. Betrachten Sie die Nutzeroberfläche gut und erkennen Sie all ihre Funktionalitäten. Suchen und finden Sie danach ein spezifiziertes Produkt. Wie Sie das Produkt finden ist Ihnen überlassen.



VERSUCH STARTEN

ZURÜCK

Abbildung C.11: Versuchsbeschreibung

Anhang D

Versuchsexperimente - Screenshots

D.1 Versuchsinstruktionen

Für jeden Versuchstest werden Versuchsinstruktionen zur Verfügung gestellt. Diese Anleitung ist als modeless Feedback implementiert und kann jederzeit geschlossen oder geöffnet werden.

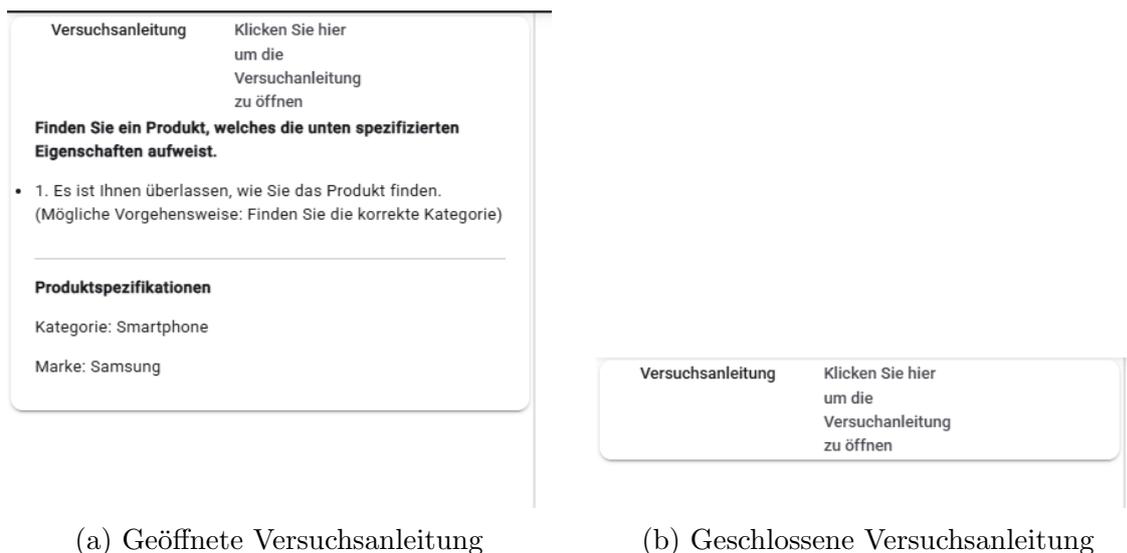
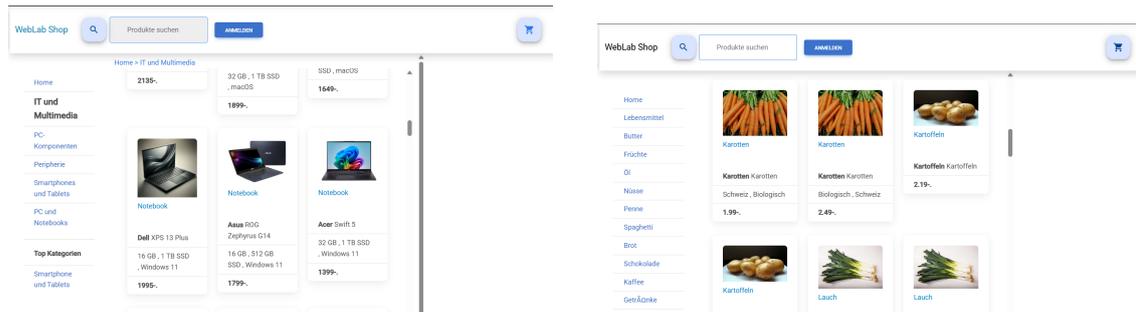


Abbildung D.1: Versuchsanleitung eines aktiven Versuches

D.2 Millers Gesetz

D.2.1 Versuchsvarianten zum Prinzip Erkennen

Die Schnittstellen der zwei Erinnerungsversuche sind analog aufgebaut. Der einzige Unterschied ist, dass die Suchfunktionen eingesetzt werden müssen.



(a) Erster Versuchstest der Kategorie Erinnern vs. Erkennen

(b) Zweiter Versuchstest der Kategorie Erinnern vs. Erkennen

Abbildung D.2: Zwei Versuchstests der Kategorie Erinnern vs. Erkennen

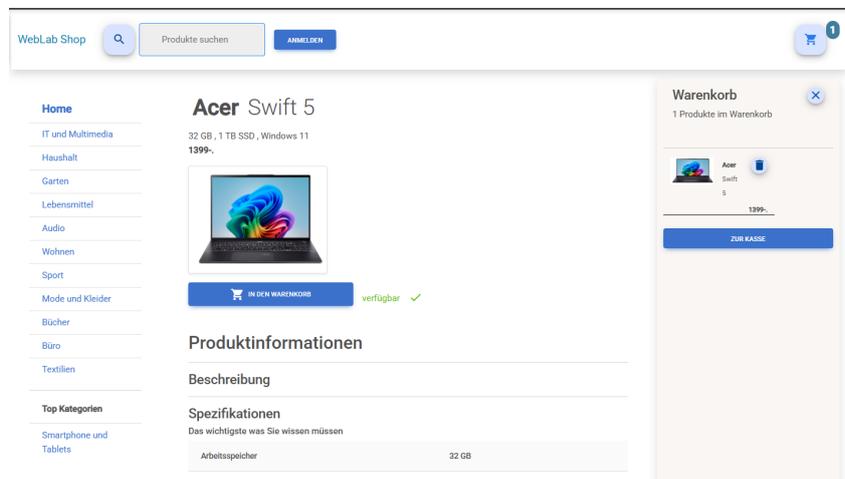


Abbildung D.3: Ein Notebook wurde ausgewählt und dem Warenkorb hinzugefügt

D.2.2 Suchfunktionen

Insgesamt wurden zwei Arten von Suchfunktionen implementiert: Eine Standard-suchfunktion ohne Autovervollständigung und eine Suchfunktion mit Autovervollständigung. Die beiden Varianten werden in Abbildung D.4 dargestellt.

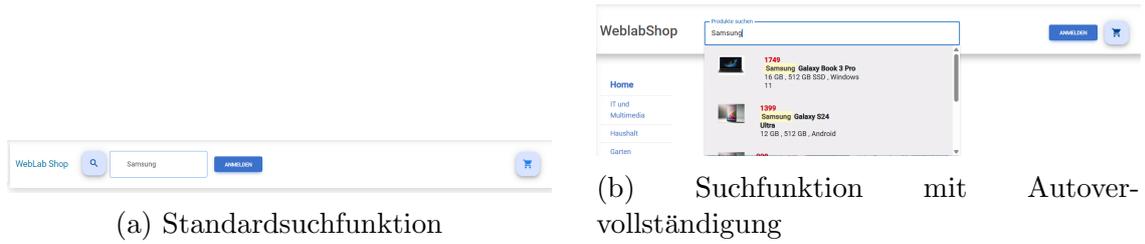


Abbildung D.4: Zwei Typen von Suchfunktionen

D.3 Gesetz von Hick

Um das Gesetz von Hick zu untersuchen, wurden drei Versuchsvarianten entwickelt.

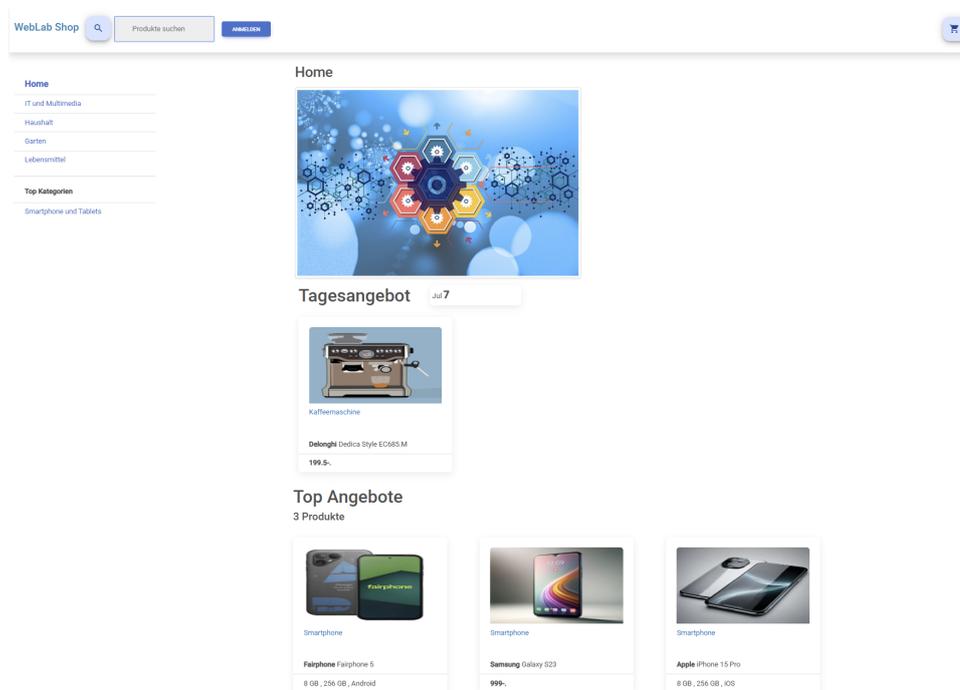


Abbildung D.5: Erster Versuch zum Gesetz von Hick mit vier Menüeinträgen

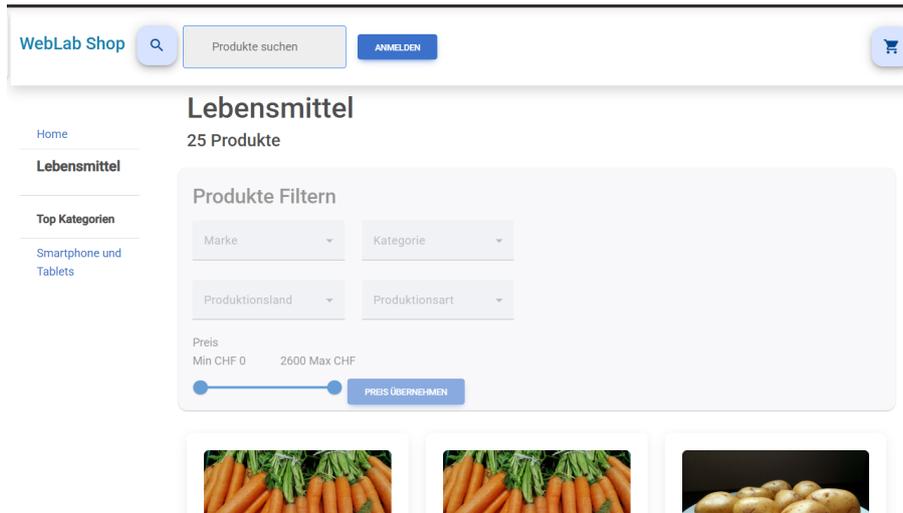


Abbildung D.6: Erster Versuch zum Gesetz von Hick: Der Link „Lebensmittel“ wurde betätigt

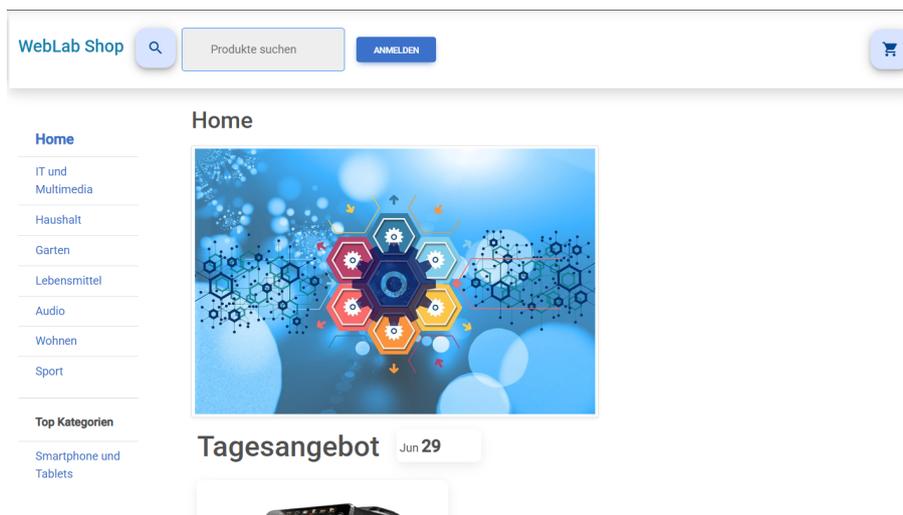


Abbildung D.7: Gesetz von Hick: Zweiter Versuchstest

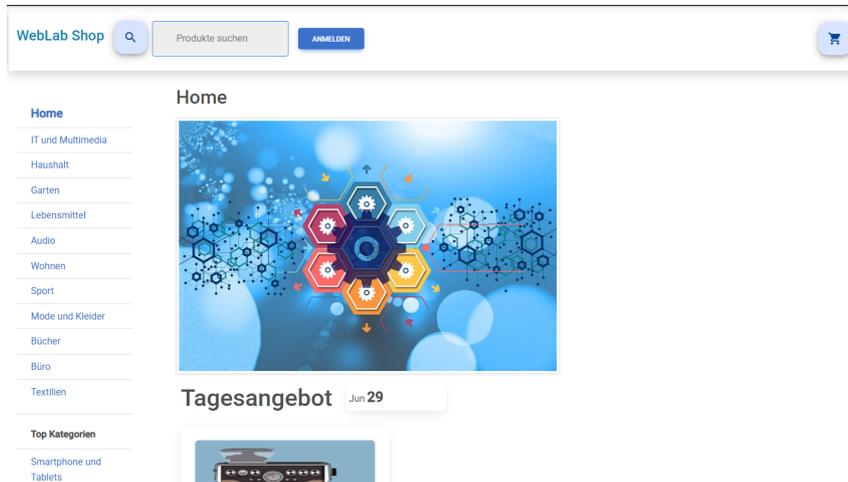


Abbildung D.8: Gesetz von Hick: Dritter Versuchstest

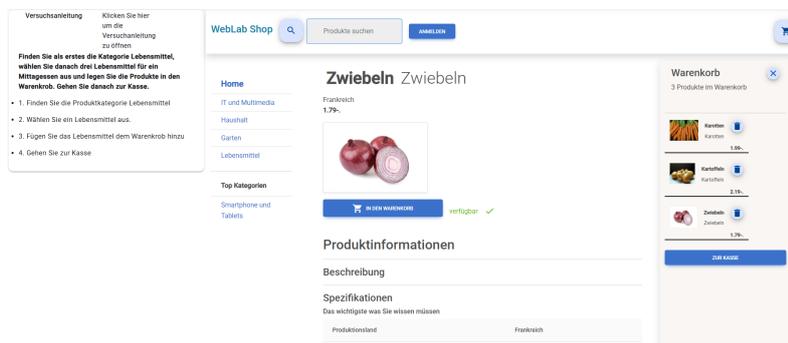
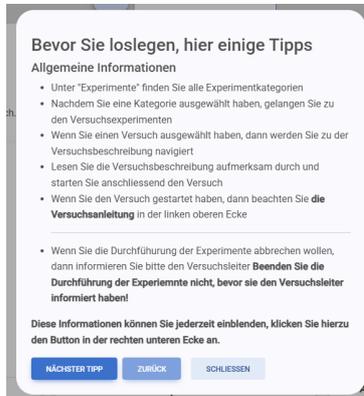


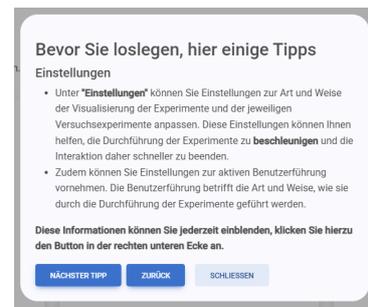
Abbildung D.9: Abschluss eines Experimentes zu Hicks Gesetz

D.4 Paradox des aktiven Nutzers

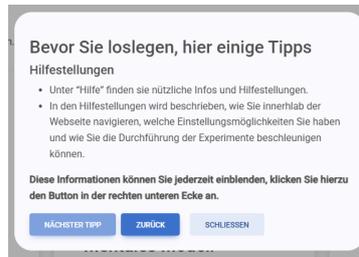
D.4.1 Dialog, der jederzeit geschlossen werden kann



(a) Erste Seite des Dialogs



(b) Zweite Seite des Dialogs



(c) Dritte Seite des Dialogs

Abbildung D.10: Willkommens-Dialog

D.4.2 Öffnen des Dialoges

Das Betätigen des blauen runden Buttons in Abbildung D.11a öffnet den Dialog, abgebildet in Abbildung D.11b

D.4.3 Dialog der Gruppe B

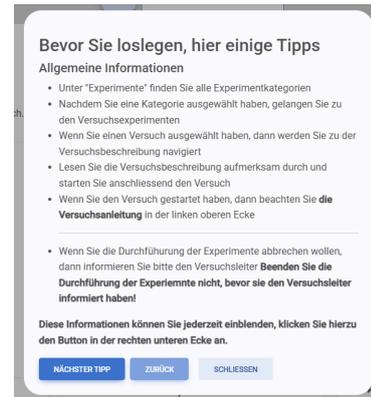
Der Dialog der Gruppe B kann erst auf der dritten Seite geschlossen werden. Der Button mit der Beschriftung „schliessen“ ist deaktiviert, bis die dritte Seite erreicht wurde.

D.4.4 Tipp-Dialog

Der in Abbildung D.13 abgebildete Dialog weist die Nutzer nach Abschluss der zweiten Kategorie auf die Einstellungsmöglichkeit der progressiven Offenlegung hin.



(a) Blauer Informationsbutton



(b) Informationsdialog nach dem Betätigen des blauen Informationsbuttons

Abbildung D.11: Öffnen des Dialoges durch den Informationsbutton links

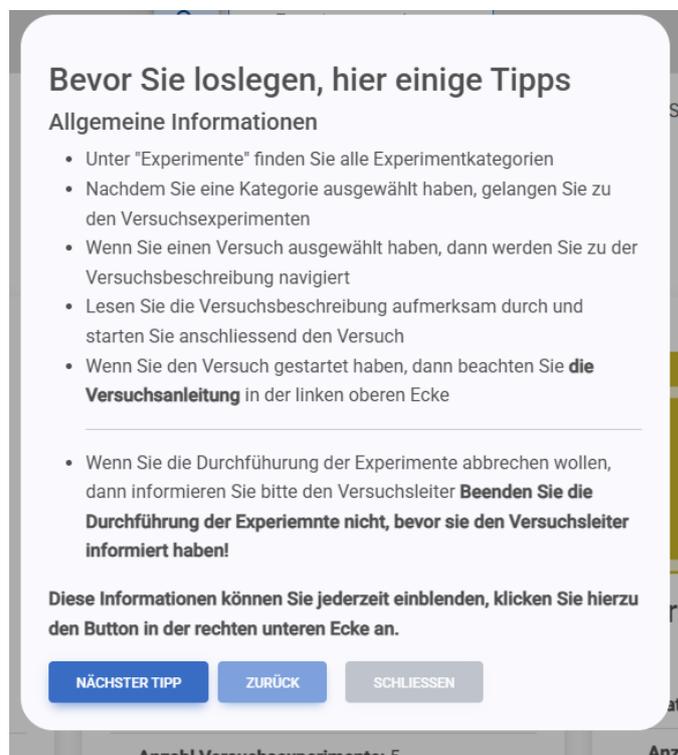


Abbildung D.12: Visualisierung des Dialoges für Gruppe B; Der Dialog kann erst auf der dritten Seite geschlossen werden

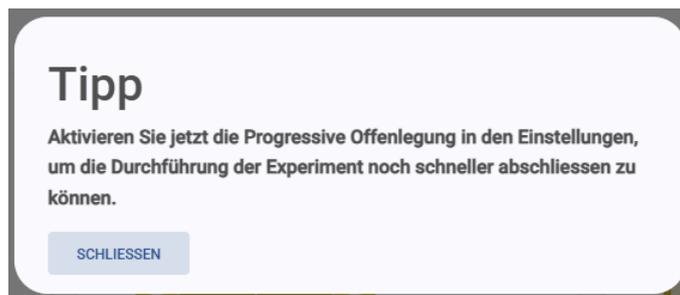


Abbildung D.13: Dialog, der die Nutzer auf die progressive Offenlegung hinweist

D.5 Teslers Gesetz

D.5.1 Progressive Offenlegung

Die Hauptseite zeigt ohne die Aktivierung der progressiven Offenlegung alle Kategorien zugleich zur Auswahl an.

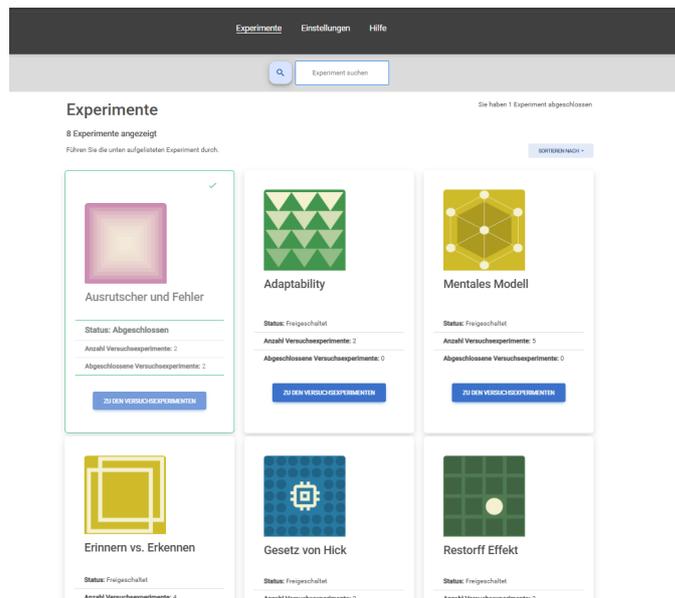


Abbildung D.14: Die Liste der Kategorien ohne die progressive Offenlegung

Wird die progressive Offenlegung aktiviert, so werden nur noch Kategorien angezeigt, welche bereits abgeschlossen wurden und jeweils eine noch nicht abgeschlossene Kategorie.

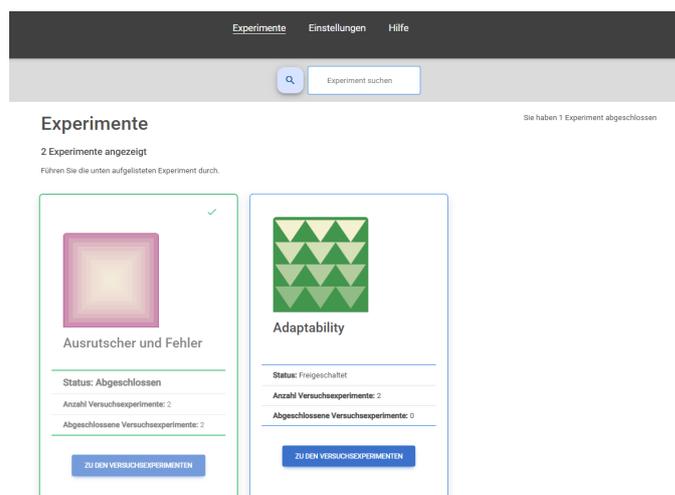


Abbildung D.15: Die Liste der Kategorien mit Aktivierung der progressiven Offenlegung

D.6 Kontrollierbarkeit und Steuerbarkeit

Für beide Versuchsvarianten zur Kontrolle und Steuerbarkeit wurde die in Abbildung D.16 dargestellte E-Mail- Anwendung eingesetzt.

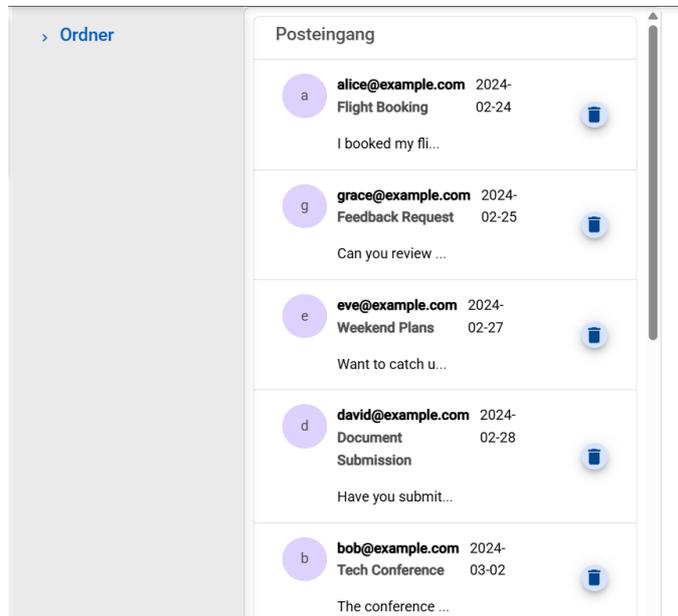


Abbildung D.16: Die grafische Nutzeroberfläche der E-Mail Anwendung

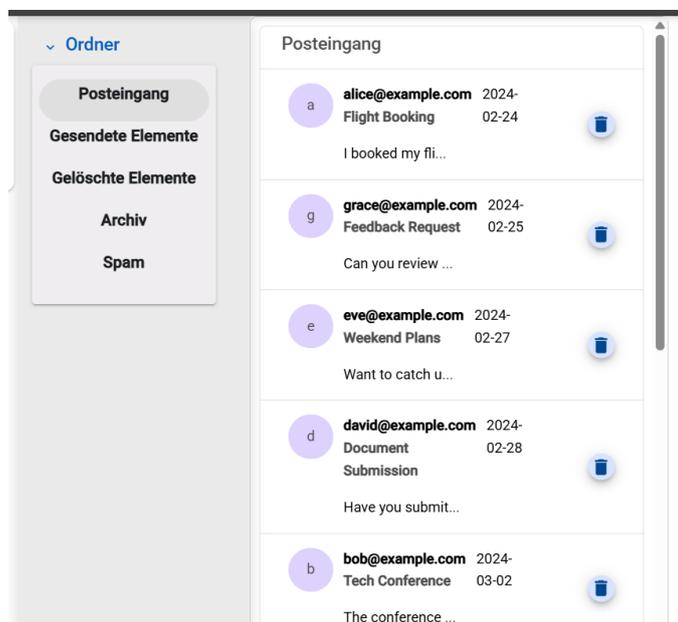


Abbildung D.17: Die grafische Nutzeroberfläche der E-Mail Anwendung mit geöffnetem Menü

Wird in der zweiten Versuchsvariante über die Undo-Funktion die E-Mail wiederhergestellt, dann wird sie in der Liste farblich hervorgehoben.

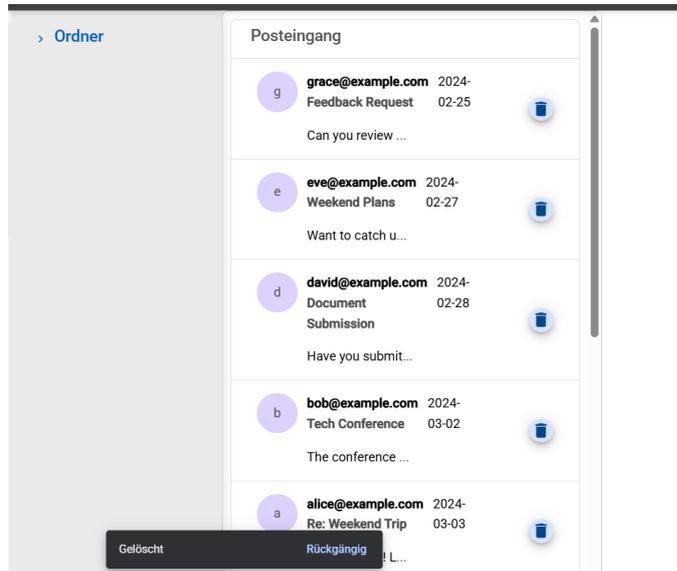
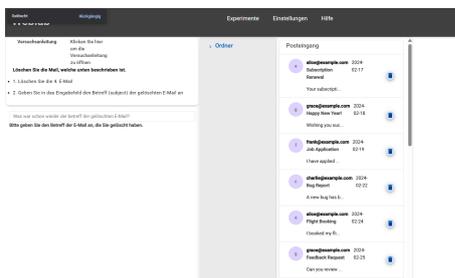
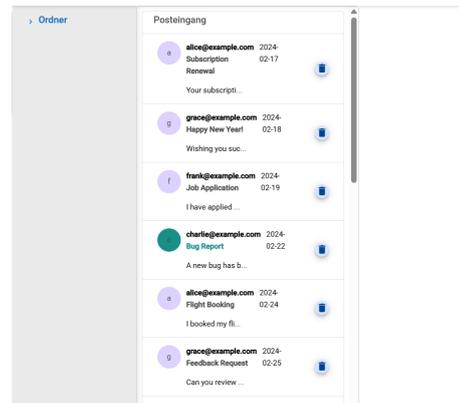


Abbildung D.18: Undo-Funktion, welche nach dem Löschen einer E-Mail für 8 Sekunden angezeigt wird



(a) Undo-Funktion in der rechten oberen Ecke



(b) Nach Betätigung der Undo-Funktion wird die E-Mail an der exakt selben Position farblich hervorgehoben eingefügt

Abbildung D.19: Wiederherstellen der gelöschten E-Mail in der zweiten Versuchsvariante

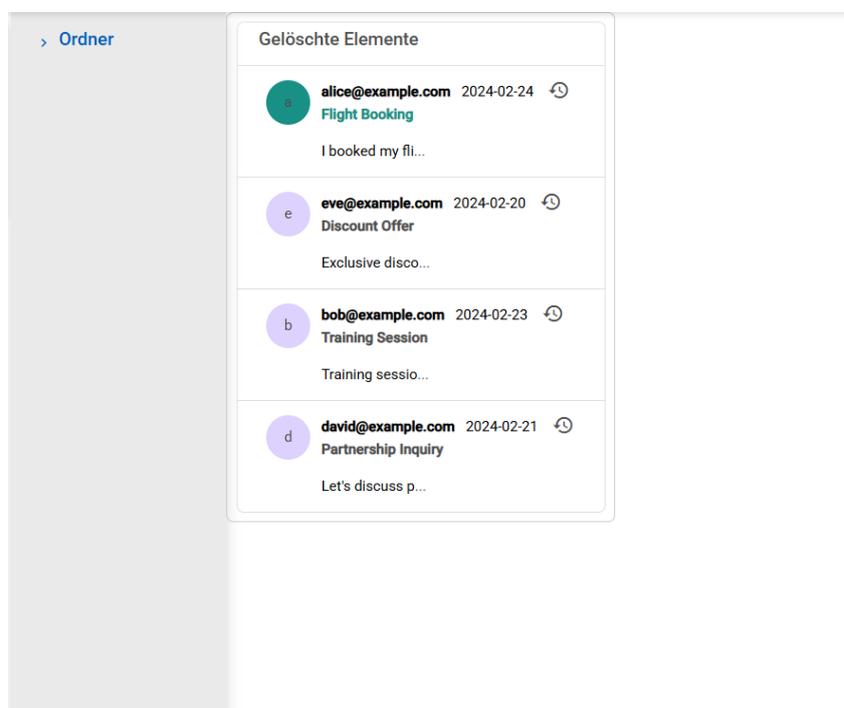


Abbildung D.20: Im zweiten Versuch wird die gelöschte E-Mail in der Liste der gelöschten Elemente farblich hervorgehoben dargestellt

D.7 Jakob's Gesetz und mentale Modelle

D.7.1 Nutzerkonfigurationen

Das Konfigurationsmenü der ersten Versuchsvariante. In dieser Illustration wurden noch keine Elemente konfiguriert. Konfigurierbare Elemente sind blau umrandet.

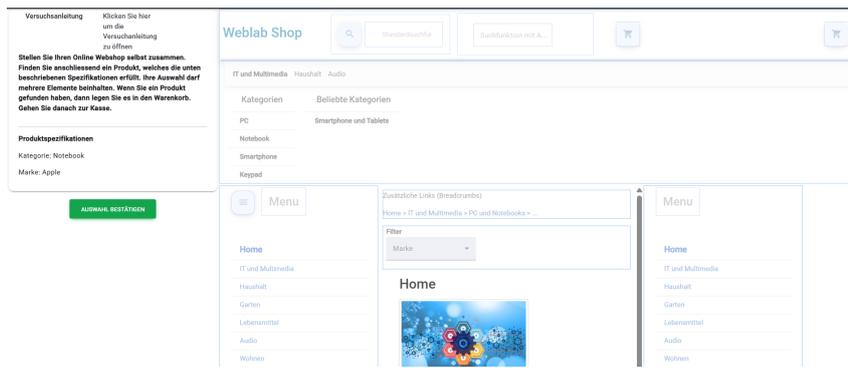


Abbildung D.21: Schnittstellen-Konfigurator

Schnittstellen-Konfigurator, der bereits konfigurierte Elemente enthält. Konfigurierte Elemente sind grün umrandet.

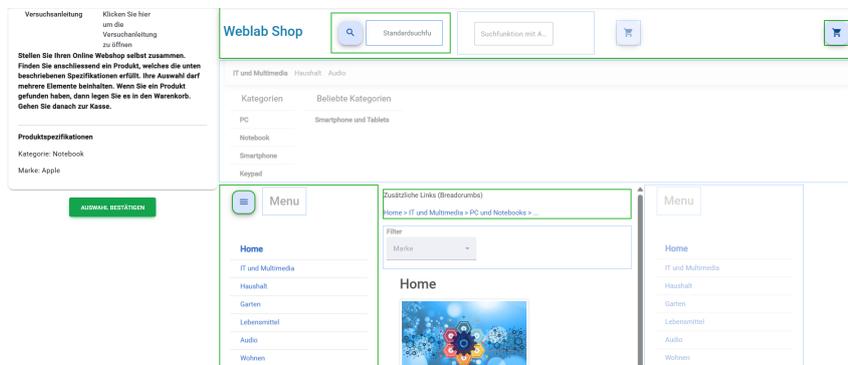


Abbildung D.22: Schnittstellen-Konfigurator mit bereits konfigurierten Elementen

In Abbildung D.22 oben wurde eine Schnittstelle konfiguriert. Abbildung D.23 zeigt nun die resultierende Schnittstelle.

Nachdem die Konfiguration abgeschlossen und bestätigt wurde, kann der Konfigurator wieder aufgerufen werden.

D.7.2 Usability-Issue

Werden der Schnittstelle Filterfunktionen hinzugefügt und die Konfiguration bestätigt, so zeigt die Schnittstelle die Filterfunktionen erst an, wenn eine Produktkategorie ausgewählt wurde.

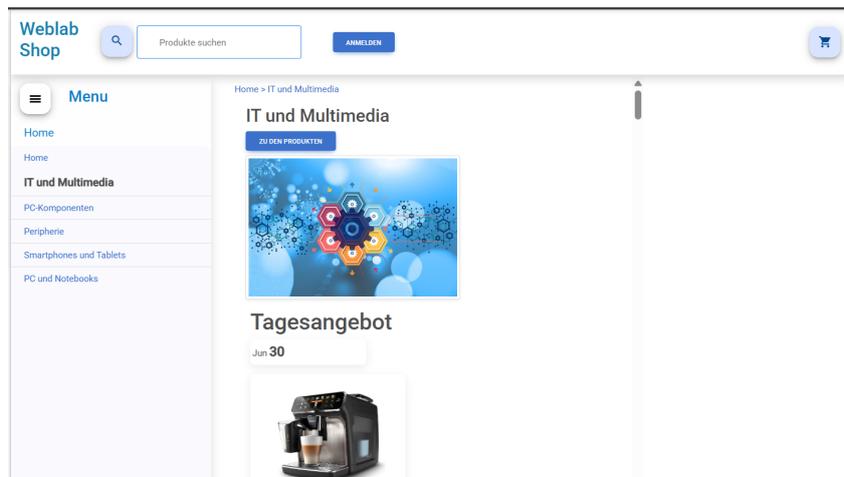


Abbildung D.23: Konfigurierte Schnittstelle

Versuchsanleitung Klicken Sie hier
um die
Versuchsanleitung
zu öffnen

**Vorher haben Sie Ihre eigene Schnittstelle
zusammengestellt. Nun müssen Sie diese verwenden um
ein Produkt zu finden, welche die unten spezifizierten
Eigenschaften erfüllt.**

- 1. Finden Sie ein Produkt, welches die folgenden
Eigenschaften aufweist und legen Sie es in den
Warenkorb.

Produktspezifikationen

Kategorie: Notebook

Marke: Apple

[ZURÜCK ZUR AUSWAHL](#)

Abbildung D.24: Nachdem die Schnittstelle konfiguriert wurde, kann zurück zum Konfigurationsmenü navigiert werden

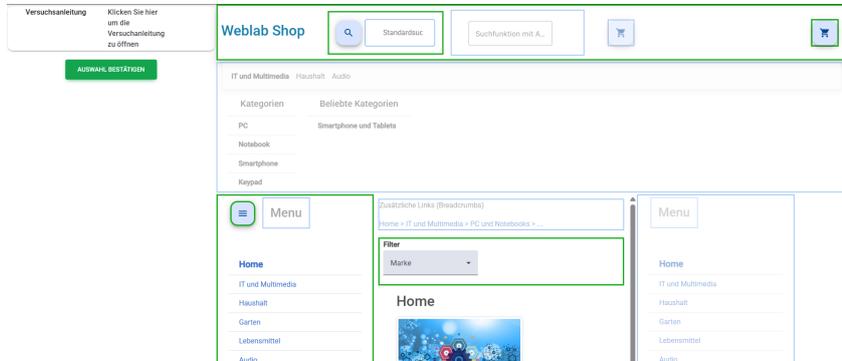


Abbildung D.25: Eine Schnittstellen wurde mit Filterfunktionen konfiguriert

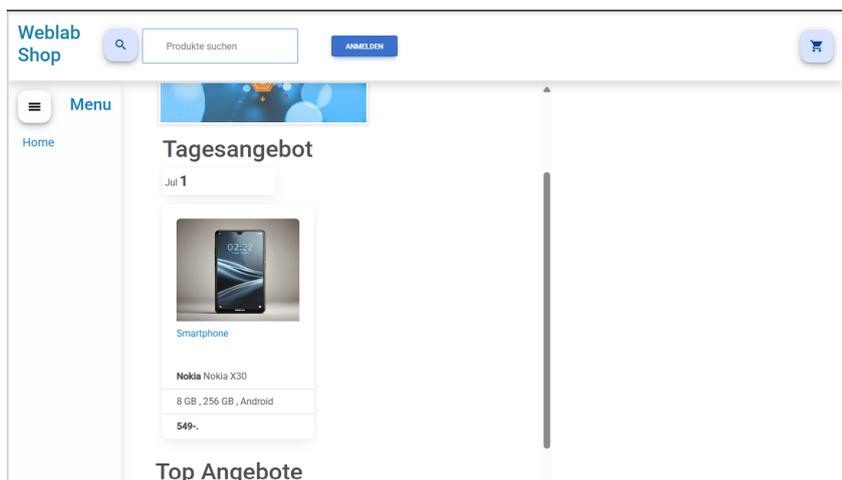


Abbildung D.26: Die Filterfunktionen werden noch nihct angezeigt, da noch keine Produktkategorie ausgewählt wurde

D.8 5 Schnittstellen zur Untersuchung mentaler Modell

Nachfolgend werden Abbildungen der fünf Schnittstellenvarianten, die der Untersuchung mentaler Modelle dienen, präsentiert.

D.8.1 Off-Canvas-Menü Schnittstellen

Schnittstelle mit einer Off-Canvas-Menükomponente und einer Suchfunktion mit Autovervollständigung.

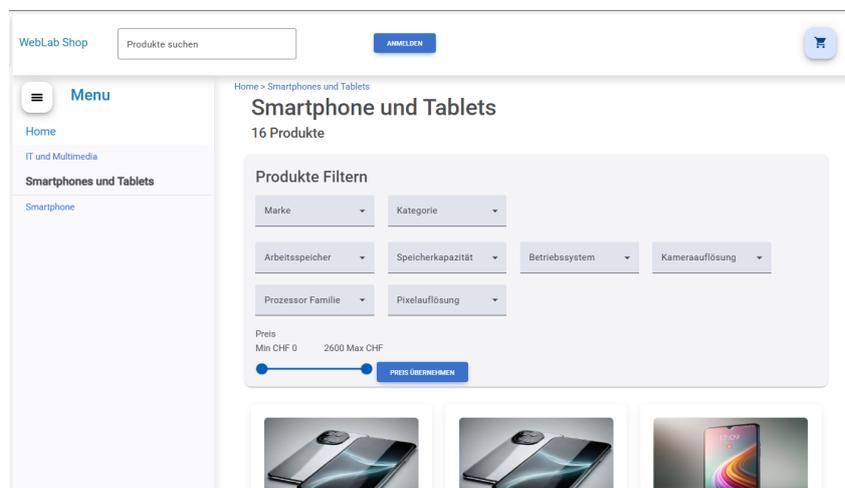


Abbildung D.27: Erste Schnittstelle zur Untersuchung mentaler Modelle mit einem Off-Canvas-Menü und einer Suchfunktion mit Autovervollständigung

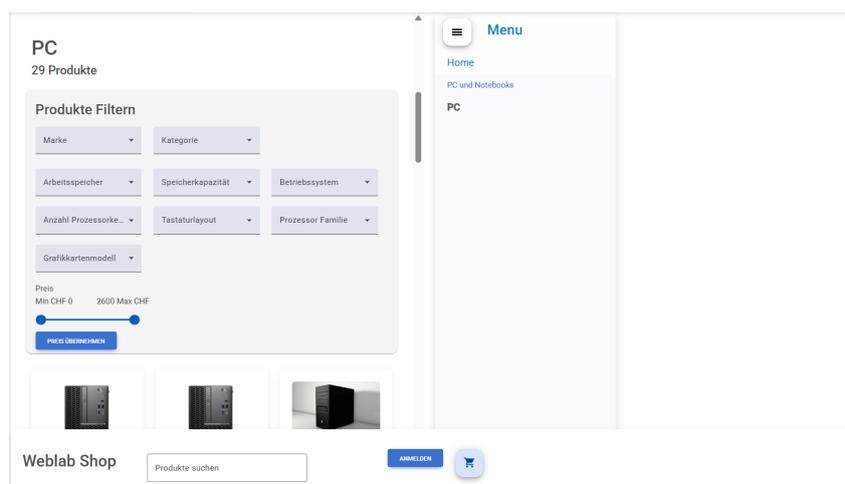


Abbildung D.28: Schnittstelle zur Untersuchung mentaler Modelle mit einem Off-Canvas-Menü am rechten Seitenrand und einer Suchfunktion am unteren Bildschirmbereich

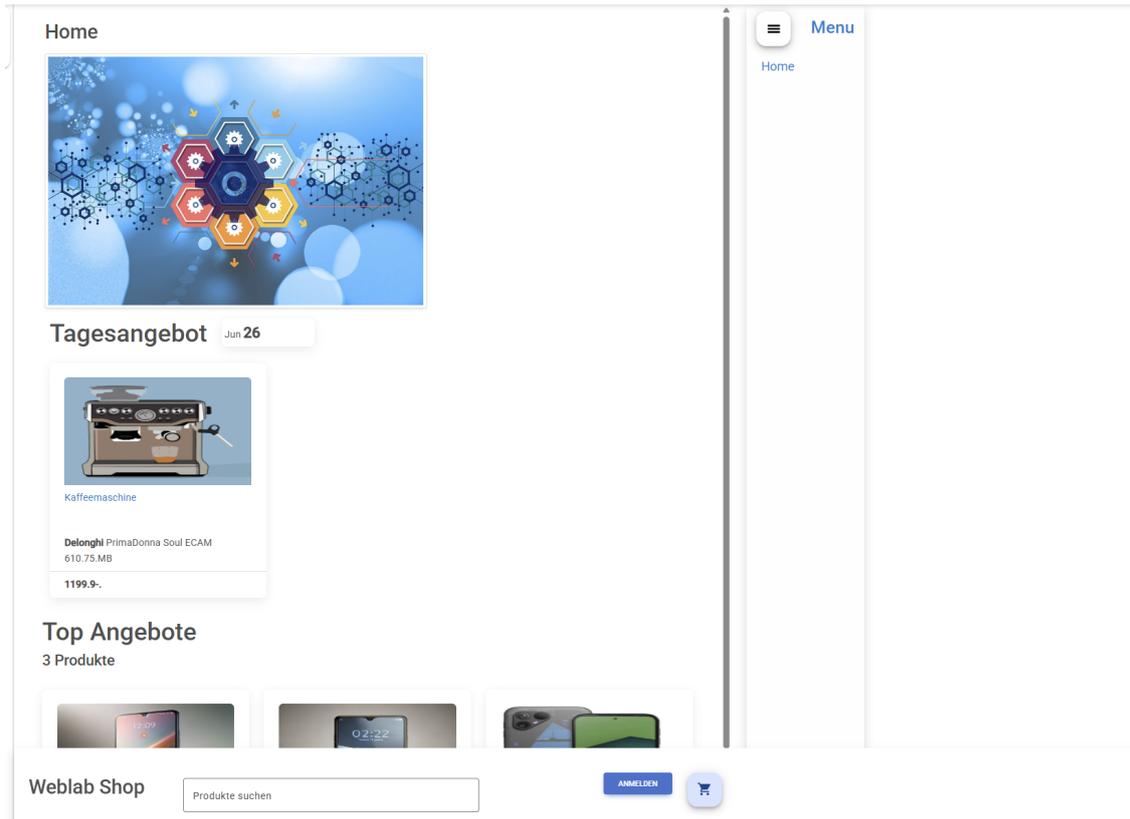


Abbildung D.29: Off-Canvas-Menü im geschlossenen Zustand

D.8.2 Mega-Drop-Down Menüschnittstelle

Schnittstelle mit einer Mega-Drop-Down-Menüschnittstelle und einer Suchfunktion ohne Autovervollständigung.

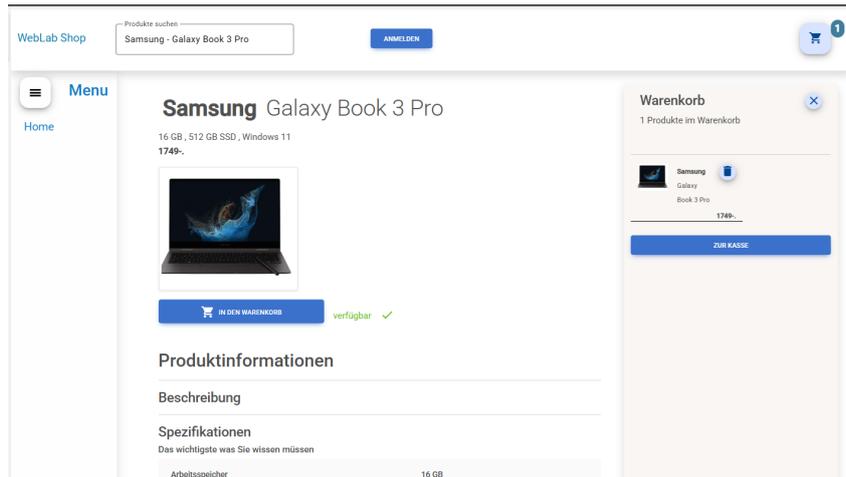


Abbildung D.30: Ein Produkt befindet sich im Warenkorb

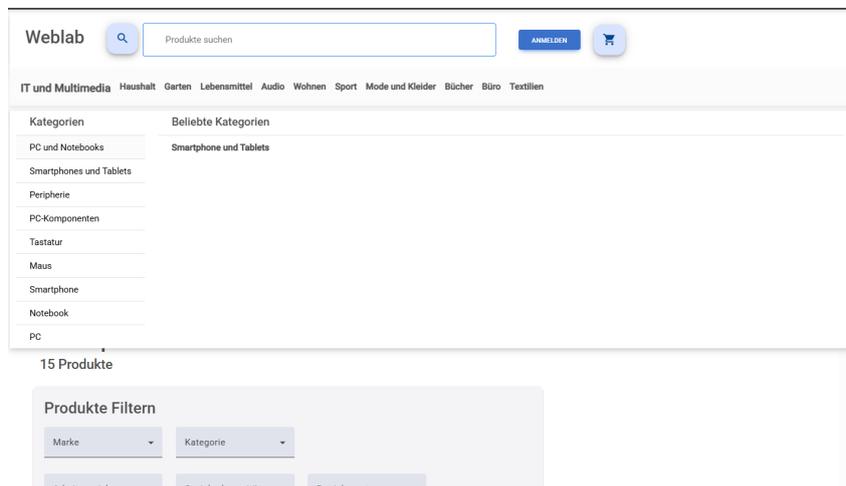


Abbildung D.31: Schnittstelle des zweiten Versuches mit einer Mega-Drop-Down Navigationsstruktur

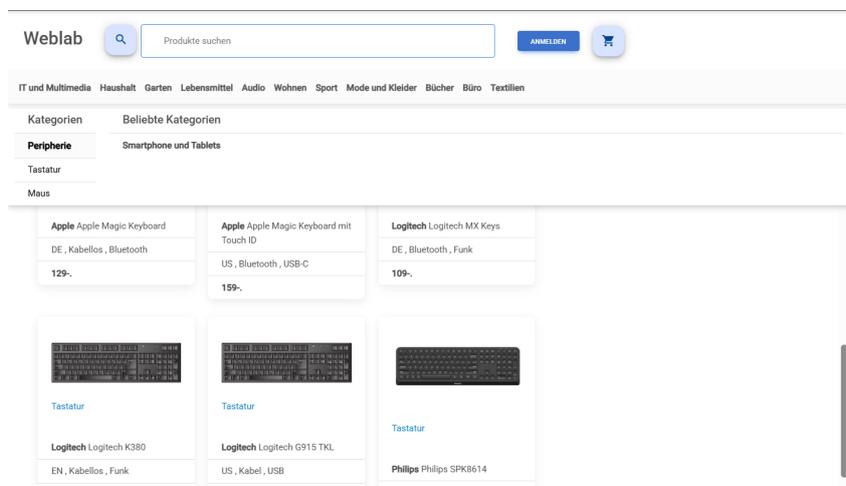


Abbildung D.32: Ein Menüeintrag wurde im Mega-Drop-Down-Menü ausgewählt

D.8.3 Schnittstellen mit textbasierten Menükomponenten (Kein Off-Canvas)

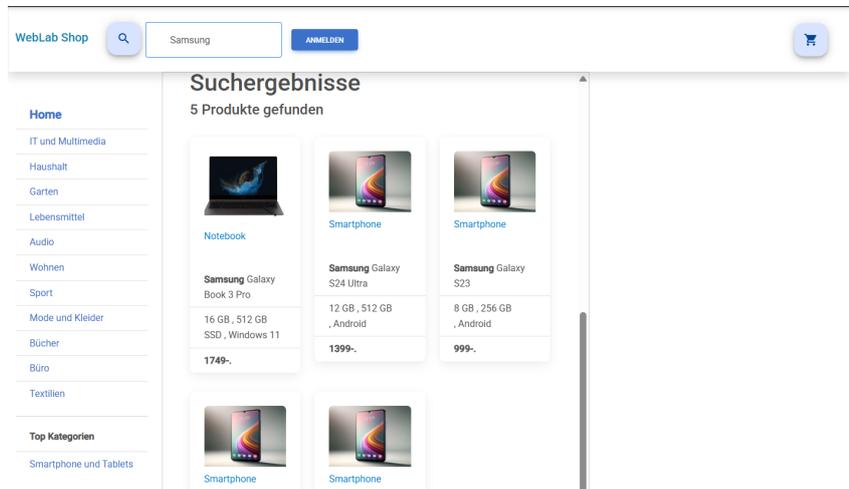


Abbildung D.33: Schnittstelle des dritten Versuches zur Untersuchung mentaler Modelle

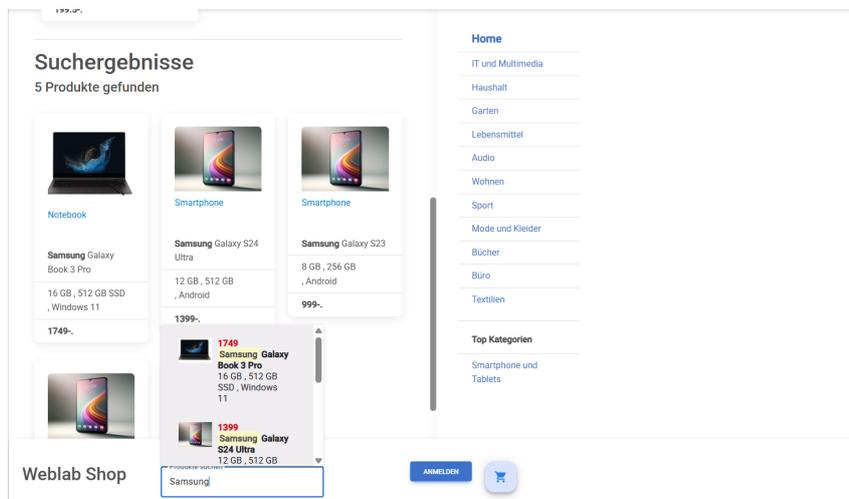
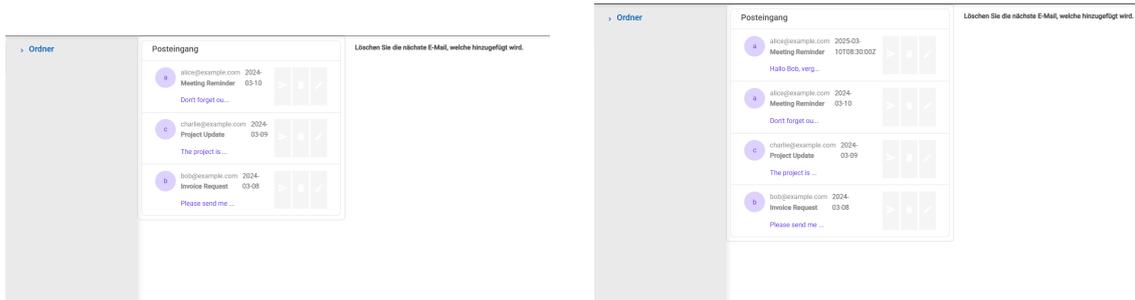


Abbildung D.34: Schnittstelle des vierten Versuches zur Untersuchung mentaler Modelle mit einem textbasierten Menü am rechten Seitenrand und einer Suchfunktion mit Autovervollständigung am unteren Bildschirmbereich

D.9 Restorff-Effekt

D.9.1 Erster Versuchstest

Im ersten Versuchstest wird der Restorff-Effekt nicht eingesetzt. Neu hinzugefügte E-Mails werden nicht farblich hervorgehoben und die Schaltflächen unterscheiden sich farblich nicht.



(a) Ausgangszustand im ersten Versuch zu Restorffs Effekt

(b) Eine E-Mail wurde der Liste hinzugefügt ohne farbliche Hervorhebung

Abbildung D.35: Funktionsweise des ersten Versuches zu Restorffs Effekt

D.9.2 Zweiter Versuchstest

Im zweiten Versuchstest wird der Restorff-Effekt leicht eingesetzt. Der Betreff neu hinzugefügter E-Mails wird farblich hervorgehoben. Die Schaltflächen unterscheiden sich farblich nicht.



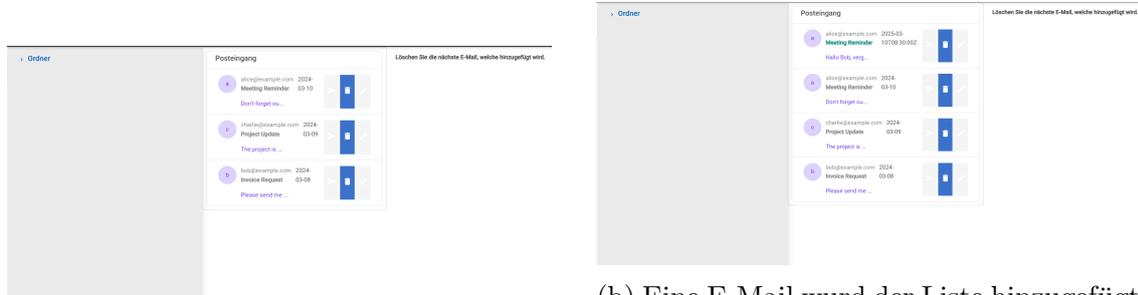
(a) Ausgangszustand im zweiten Versuch zu Restorffs Effekt

(b) Eine E-Mail wurde der Liste hinzugefügt mit farblicher Hervorhebung des E-mail Betreffs

Abbildung D.36: Funktionsweise des zweiten Versuches zu Restorffs Effekt

D.9.3 Dritter Versuchstest

In der dritten Versuchsvariante unterscheidet sich der Lösch-Button farblich von den anderen Aktionselementen.



(a) Ausgangszustand im dritten Versuch zu Restorffs Effekt

(b) Eine E-Mail wurde der Liste hinzugefügt mit farblicher Hervorhebung des E-mail Betreffs

Abbildung D.37: Funktionsweise des dritten Versuches zu Restorffs Effekt

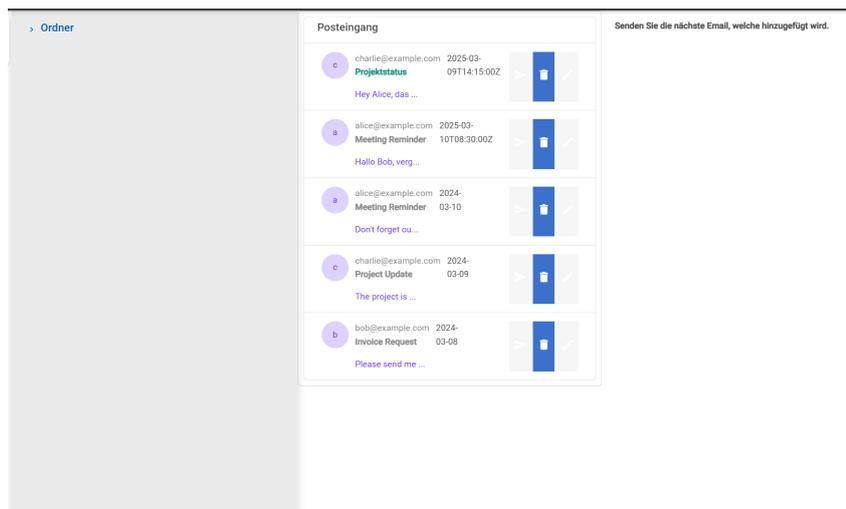


Abbildung D.38: Die dritte und fünfte E-Mail muss anders als die restlichen jeweils weitergeleitet werden. Die Instruktion befindet sich rechts von der Liste

Literaturverzeichnis

- [1] Michaela Kauer-Franz and Benjamin Franz. *Usability und User Experience Design*. Number 1. Rheinwerk Verlag, 1 edition.
- [2] Jens Jacobsen and Lorena Meyer. *Usability und UX, Praxisbuch*. Number 2. Rheinwerk Verlag, 4 edition.
- [3] M. Andreas Heinecke and Jens Gerken. *Mensch-Computer-Interaktion*. Number 4. Springer Fachmedien, 3 edition.
- [4] Jon Yablonski. *Laws of UX, 10 praktische Grundprinzipien für intuitives, menschenzentriertes UX-Design*. Number 3. dpunkt.verlag, 2 edition.
- [5] I. Scott MacKenzie. *Human-Computer Interaction, An Empirical Research Perspective*. Number 5. Morgan Kaufmann, 2 edition.
- [6] John Carroll and Mary Beth Rosson. Paradox of the active user in interfacing thought: cognitive aspects of human computer interaction. *MIT Press*, pages 80–111, 01 1987.
- [7] Monika Heimann and Michael Schütz. Wie design wirkt, psychologische prinzipien erfolgreicher gestaltung.
- [8] George Miller. The magical number seven, plus or minus two: Some limits on out capacity for processing information. *Psychological review*, 101(6):343–52, 04 1994.
- [9] John Sweller. The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review*, 35(7), 09 2023.
- [10] Barry Schwartz. *The Paradox of Choice: Why More is Less*. HarperCollins.
- [11] Radka Nacheva. The importance of users' mental models for developing usable human- machine interfaces. 10 2015.

- [12] Hosam Al-Samarraie, Hassan Selim, Timothy Teo, and Fahed Zaqout. Isolation and distinctiveness in the design of elearning systems influence user preferences. *Interactive Learning Environments*, 25:1, 02 2016.
- [13] Andrea Marrella and Tiziana Catarci. Measuring the learnability of interactive systems using a petri net based approach.
- [14] Pradip Peter Dey, Bhaskar Raj Sinha, Mohammad Amin, and Hassan Badkoo-behi. Best practices for improving user interface design. *International Journal of Software Engineering AND Applications*, 10(5).
- [15] Rajiv Choudhary and Prasun Dewan. A general multi-user undo/redo model. *Synthesis Lectures on Human-Centered Informatics*, 08 1998.
- [16] Loredana Provenza and Antonio Piccinno. Designing flexible user interfaces. (11):539–547, 01 2010.
- [17] Chakradhar Devarapalli. Accessibility in user interfaces: Confronting common challenges. *Journal of Artificial Intelligence and Cloud Computing*, pages 1–4, 06 2022.
- [18] Thomas Tullis. Designing a menu-based interface to an operating system. *Proceedings of the of the CHI' 85 conference on human factors in computing systems*, 16:79–84, 04 1985.
- [19] Jeremy Mendel and Richard Pak. The effect of interface consistency and cognitive load on user performance in an information search task. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 53(12):1684–1688, 10 2009.
- [20] Don Norman. *The Design of Everyday Things*. Basic Books, revised and exappnded edition, 2013.
- [21] B. Shneidermann and C. Plaisant. Designing the user interface: Strategies for effective human-computer interaction. (4).
- [22] Bill Albert and Tom Tullis. *Measuring the user experience, Collecting, analyzing, and presenting ux-metrics*. Morgen Kaufmann, 3 edition.
- [23] Jeff Sauro and James R. Lewis. *Quantifying the user Experience*. Morgan Kaufmann, 2 edition.

- [24] Morten Hertzum. Usability testing: A practitioner's guide to evaluating the user experience. *Synthesis Lectures on Human-Centered Informatics*, 1:i–105, 03 2020.
- [25] Dominik Mayer, Martin Schrepp, and Theo Held. Beurteilung der ux qualität durch experten. *Mensch und Computer 2018 - UX Professionals*, (1):3–14.
- [26] James Lewis. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: The asq. *SIGCHI Bull.*, 23:78–81, 01 1991.
- [27] J. D. Brroke. SUS: A quick and dirty usability scale. in p. w. jorand b. thomas, b. a. weerdmeester, & i. l. McCelland (eds.). *usability evaluation in industry*.
- [28] Martin Schrepp. User experience questionnaire handbook, 09 2015.
- [29] P. A. Smith. Towards a practical measure of hypertext usability. *Interacting with Computers*, 4(8):365–381, 1996.
- [30] Wang Junfeng, Zhiyu Xu, Xi Wang, and Jingjing Lu. A comparative research on usability and user experience of user interface design software. *International Journal of Advanced Computer Science and Applications*, 13, 01 2022.

Erklärung

gemäss Art. 30 RSL Phil.-nat.18

Name/Vorname: Brentani Romano

Matrikelnummer: 21-103-668

Studiengang: Informatik

Bachelor

Master

Dissertation

Titel der Arbeit: Die Wirkung von UX-Gesetzen auf die Usability und User Experience interaktiver Webschnittstellen

LeiterIn der Arbeit: PD Dr. Kaspar Riesen

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe r des Gesetzes vom 5. September 1996 über die Universität zum Entzug des auf Grund dieser Arbeit verliehenen Titels berechtigt ist. Für die Zwecke der Begutachtung und der Überprüfung der Einhaltung der Selbständigkeitserklärung bzw. der Reglemente betreffend Plagiate erteile ich der Universität Bern das Recht, die dazu erforderlichen Personendaten zu bearbeiten und Nutzungshandlungen vorzunehmen, insbesondere die schriftliche Arbeit zu vervielfältigen und dauerhaft in einer Datenbank zu speichern sowie diese zur Überprüfung von Arbeiten Dritter zu verwenden oder hierzu zur Verfügung zu stellen.

Bern, 11.08.2025

Ort/Datum

Unterschrift

R. Brentani