

Simulation des Interaktionsverhaltens von Senioren bei der Benutzung von mobilen Endgeräten

vorgelegt von
Dipl.-Inf., Dipl.-Desig. Electronic Business
Matthias Schulz
geb. in Königs Wusterhausen

von der Fakultät IV – Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
– Dr.-Ing. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr.-Ing. Olaf Hellwich
Gutachter: Prof. Dr.-Ing. Sebastian Möller
Gutachter: Prof. Dr. Hartmut Wandke
Gutachter: Dr. Sebastian Feuerstack

Tag der wissenschaftlichen Aussprache: 11.12.2015

Berlin 2016

Matthias Schulz: *Simulation des Interaktionsverhaltens von Senioren bei der Benutzung von mobilen Endgeräten.* © 30. Juni 2015

KURZDARSTELLUNG

Das Ziel dieser Arbeit ist es, eine bestehende Software zur Automatischen Usability Evaluierung (AUE) dahingehend zu erweitern, dass demografische Verteilungen über Einschränkungen in der älteren Bevölkerung in den Simulationsprozess einfließen und Auswirkungen von fehlenden oder unpassenden mentalen Modellen des Eingabegerätes simuliert werden können. Die vorliegende Arbeit dokumentiert den aktuellen Stand der Forschung, die durchgeführten empirischen Versuche zur Modellbildung bzw. Validierung der Modelle und die Weiterentwicklung der Software.

Wie in dieser Arbeit gezeigt wird, haben ältere Menschen verschiedene Einschränkungen (Wahrnehmung, Kognition und Motorik), die die Interaktion mit einem System negativ beeinflussen können. Es existieren verschiedene Methoden und Werkzeuge, die es ermöglichen Usability-Fehler durch Simulation zu finden, allerdings haben die Methoden und Werkzeuge offene Probleme, die die Anwendbarkeit limitieren. So wird nicht betrachtet, ob die Personen mit dem Eingabegerät zurechtkommen. Weiterhin werden nur wenige oder keine demografischen Daten verwendet, die eine Abschätzung der Schwere eines Usability-Problems ermöglichen könnten.

Um die demografischen Verteilungen von Einschränkungen in die AUE integrieren zu können, wird eine technische Lösung beschrieben, die es ermöglicht Nutzermodelle zu generieren, die realistische Einschränkungen haben. Durch die Reproduktion der Verteilungen von Einschränkungen in der Gesellschaft ist es möglich die Schwere eines Usability-Problems abzuschätzen. Zur Integration von unpassenden und fehlenden mentalen Modellen von Eingabegeräten in die Simulation wurden zwei Versuche durchgeführt. Der erste Versuch diente der Modellbildung und der zweite Versuch der Validierung der Modelle durch einen unabhängigen Test. Die Ergebnisse beide Versuche werden in dieser Arbeit dargestellt.

Durch die korrekte Reproduktion von Einschränkungen in der Gesellschaft und durch die beispielhafte Einbindung von mentalen Modellen bezüglich von Eingabegeräten, kann die Qualität von Simulationen substanziell gesteigert werden. Die in dieser Arbeit vorgestellten Ergebnisse erweitern den aktuellen Stand der Forschung und ermöglichen eine vielseitige Weiternutzung.

ABSTRACT

This thesis focuses on the simulation of interaction behaviour of older adults using input devices which are operated by a touchscreen (e. g. smartphones or tablets). The simulation is used to predict usability problems.

One goal of this thesis is to enhance a tool for Automated Usability Evaluation (AUE) with demographic distributions of cognitive and physical impairments. The second goal aims to predict interaction errors, which are caused by inappropriate mental models of how to use the input device, and to integrate the prediction into the AUE tool.

Older adults often suffer from age-related diseases which affect the perception, cognition and motor function. These limitations usually influence the human machine interaction negatively and reduce the usability of a system. There are different methods and tools which can find usability problems by simulation. However, these tools work to a limited and unsatisfactory extent, because they do not consider problems caused by missing knowledge about input device handling. Furthermore, they only use minor or no demographic data within the simulation, which makes the estimation of a problem difficult and unreliable.

A part of this thesis is the description of a technical implementation, which brings demographic data into a tool for AUE simulation. The technical implementation allows the generation of user models having realistic descriptions of limitations and diseases. By reproducing demographic distributions, it is possible to estimate the severity of a problem.

To integrate the effects of inappropriate mental models into the AUE simulation two studies are conducted. The first study was used to obtain several generalized linear models, which are used to simulate the decision process of the mental model. The second study was used to validate the linear models and the implementation. The results of both studies are presented in this thesis.

The reproduction of demographic disease distributions and the simulation of inappropriate mental models of input devices increase the value of the AUE simulation substantially.

DANKSAGUNG

Mein besonderer Dank gilt Herrn Prof. Dr.-Ing. Sebastian Möller für die umfassende Betreuung, die vielen hilfreichen Anmerkungen und seine Motivation während des Schreibens der Dissertation. Weiterhin danke ich herzlich Herrn Prof. Dr. sc. nat. Hartmut Wandke und Herrn Dr. Sebastian Feuerstack für die Anfertigung des Zweit- bzw. Drittgutachtens.

Diese Arbeit wäre ohne die Unterstützung so vieler anderer Personen nicht möglich gewesen. Ich möchte mich zuerst bei Dr.-Ing. Klaus-Peter Engelbrecht und Dr. Benjamin Weiss bedanken, die mich während meiner Promotion unterstützen, mir mehrmals den richtigen Weg zeigten und in den entscheidenden Situationen die Augen für alternative Herangehensweisen öffneten. Mein ganz besonderer Dank gilt Stefan Hillmann, der mir half Licht in das Dunkel des MeMo-Codes zu bringen und der durch seine intelligent gestellten Fragen einen großen Betrag zum Gelingen der Arbeit beitrug.

Ebenfalls einen großen Anteil zum Gelingen der Arbeit haben meine Kollegen und Freunde geleistet. Ich möchte mich hiermit bei Aaron Russ, Christine Kühnel, Dr. Hagen Wierstorf, Dr.-Ing. Ina Wechsung, Janto Skowronek, Dr. Julia Niemann, Marc Halbrügge und Dr.-Ing. Michael Quade für die lebhaften Diskussionen, interessanten Anregungen und Hilfestellungen bedanken.

Schließlich danke ich meinen Eltern, die mich immer unterstützt haben, und meiner Freundin Katrin, die mir den notwendigen Rückhalt und die Liebe gab die ich brauchte, um diese Arbeit abschließen zu können.

INHALTSVERZEICHNIS

1	EINFÜHRUNG	1
1.1	Was ist Usability?	1
1.2	Methoden zur Evaluation der Usability	3
1.3	Ältere Versuchspersonen	5
1.3.1	Begriffsdefinition	5
1.3.2	Einschränkungen	6
1.3.3	Mentale Modelle	8
1.3.4	Usability-Tests mit älteren Versuchspersonen	9
1.4	Automatische Evaluierung der Usability	11
1.4.1	Methoden und Werkzeuge	12
1.4.2	Nutzermodelle	17
1.5	Ziele und Aufbau der Arbeit	22
2	MEMO-WERKBANK	25
2.1	Teile der MeMo-Werkbank	25
2.1.1	Wahrscheinlichkeiten	26
2.1.2	Regeln und Regelverarbeitung	28
2.1.3	Systemmodell	30
2.1.4	Aufgabenmodell	35
2.1.5	Nutzermodell	36
2.2	Simulation	41
2.2.1	Ablauf	42
2.2.2	Ergebnisse	44
2.3	Fazit	44
3	DEMOGRAFISCHE NUTZERMODELLE	47
3.1	Nutzerattribute	47
3.1.1	Hören	48
3.1.2	Sehen	52
3.1.3	Kognition	56
3.1.4	Motorik	61
3.1.5	Weitere Attribute	67
3.2	Modellierung von Nutzerattributen	71
3.2.1	Einführung in Bayes'sche Netze	71
3.2.2	Repräsentation der Nutzerattribute	78
3.3	Generierung von Nutzermodellen	93
3.3.1	Ablauf	93
3.3.2	Beispiele und Validierung	94
3.4	Fazit	97
3.4.1	Positive Aspekte	98
3.4.2	Negative Aspekte	98

4	SIMULATION VON MENTALEN MODELLEN	101
4.1	Empirischer Versuch zur Modellbildung	101
4.1.1	Ziele	102
4.1.2	Aufbau und Geräte	102
4.1.3	Versuchspersonen	104
4.1.4	Ablaufplan	105
4.1.5	Aufgabenstellungen	106
4.1.6	Videoannotation der Smartphone-Interaktion	107
4.1.7	Ergebnisse	110
4.1.8	Zusammenfassung und Ableitungen zur Modellbildung	120
4.2	Erweiterung der MeMo-Werkbank	121
4.2.1	Generische Interaktionen und Device Models	122
4.2.2	Systemfeedback und neue Nutzerattribute	124
4.2.3	Modul zur Wahrnehmung der Benutzerschnittstelle	125
4.2.4	Modul zur Verarbeitung von wahrgenommenen Ereignissen	127
4.2.5	Modul zur Manipulation der Benutzerschnittstelle	131
4.2.6	Modul zur Verarbeitung der Nutzereingabe	133
4.2.7	Modul zum Export der Simulationsergebnisse	135
4.3	Simulation	136
4.3.1	Aufbau des Modells und Bedingungen für die Simulation	136
4.3.2	Aufgabenwissen der Nutzermodelle	138
4.3.3	Vergleich der empirischen Ergebnisse mit den Simulationsergebnissen	139
4.4	Fazit	151
4.4.1	Positive Aspekte	152
4.4.2	Negative Aspekte	152
5	VALIDIERUNGSVERSUCH	155
5.1	Zielstellung	155
5.2	Aufbau und Geräte	155
5.3	Versuchspersonen	158
5.4	Aufgabenstellungen	159
5.5	Ablauf	161
5.6	Ergebnisse	163
5.6.1	Empirischer Versuch	164
5.6.2	Simulation	182
5.6.3	Kombination der Ansätze	196
5.7	Fazit	205
5.7.1	Positive Aspekte	205
5.7.2	Negative Aspekte	206
6	ZUSAMMENFASSUNG UND AUSBLICK	209

6.1	Zusammenfassung	209
6.1.1	Umgesetzte Arbeiten	210
6.1.2	Erreichte Ziele	210
6.1.3	Kritische Betrachtung der erreichten Ziele	211
6.2	Ausblick	214
6.2.1	Alternative Verwendungsmöglichkeiten	214
6.2.2	Offene Arbeiten	215
A	ALLGEMEINE LINEARE MODELLE	217
A.1	Kapazitiver Touchscreen	217
A.2	Resistiver Touchscreen	217
A.3	Fernbedienung	218
B	BEISPIEL FÜR DEN EXPORT DER SIMULATIONSERGEBNISSE	221
B.1	Rohdaten	221
B.2	Zusammengefassten Daten	223
	LITERATUR	229

AKRONYME

ACT-R	Adaptive Control of Thought-Rational
AMD	Altersabhängige MakulaDegeneration
ANOVA	ANalysis Of VAriance
API	Application Programming Interface
AUE	Automatische Usability Evaluierung
BN	Bayes'sches Netz
BS	Biswas-Simulator
BUPM	Bottom-Up Perception Module
CAPS	Cognitive-Affective Processing System
CORE	Constraint-based Optimal Reasoning Engine
CT	CogTool
CT-E	CogTool-Explorer
dB	Dezibel
DEM	Default Execution Module
DIAS	Disability Impairment Approximation Simulator
DISCO	extracting DIstributionally related words using CO-occurrences
DMM	Device Mental Model
DPERM	Default PERception Module
DPROM	Default PROcessing Module
DSM	Dual-Space Model
EPIC	Executive-Process/Interactive Control
GeNIe	Graphical Network Interface
GOMS	Goals, Operators, Methods, and Selection rules
HCI	Human Computer Interaction
Hz	Hertz

ICT	Information and Communication Technology
IQ	Intelligenzquotient
JSON	JavaScript Object Notation
HADRIAN	Human Anthropometric Data Requirements Investigation and ANalysis
INSPIRE	INfotainment management with SPeech Interaction via REMote microphones and telephone interfaces
KLM	Keystroke-Level Model
MASP	Multi-Access Service Platform
MeMo	Mental models
MHP	Model Human Processor
MPM	Multimodal Processing Module
MS	Multiple Sklerose
PPM	Payment Processing Module
SDS	Sprachdialogsystem
SHT	Schädel-Hirn-Trauma
SMILE	Structural Modeling, Inference, and Learning Engine
SMM	System Mental Model
SNIF-ACT	Scent-based Navigation and Information Foraging in the ACT architecture
SOAR	State, Operator Apply Result
SP	Sample Space
SPM	Semantic Processing Module
SSM	System Mental Model
TA	Technikaffinität
TU	Technische Universität
UPM	Utterance Processing Module
XML	eXtensible Markup Language

FREMDWÖRTER

A-posteriori	Abhängig bzw. aus der Erfahrung abgeleitet
A-priori	Unabhängig bzw. aus der Vernunft abgeleitet
Adipositas	Starkes Übergewicht
Anthropometrisch	exakt und wissenschaftlich gehobene Körpermerkmale
Aphasie	Sprachstörung
App	Programm auf einem Smartphone
Azyklisch	Ohne Zyklen
Cochlea	Hörschnecke
Evidenz	Beobachtete Fakten, Wissen
Fluide Intelligenz	Fähigkeit die wahrgenommenen Informationen schnell und akkurat zu verarbeiten
Fovea centralis	Bereich des schärfsten Sehens
Glaukom	Schädigung des Sehnerves, führt meist zu Gesichtsfeldausfällen
Homoskedastizität	Varianz der Verteilung
Hyperopie	Weitsichtigkeit
Hypertonie	Bluthochdruck
Katarakt	Meist altersbedingte Trübung der Augenlinse
Kognition	Intelligente Verarbeitung von wahrgenommenen Informationen durch das Gehirn
Kristalline Intelligenz	Durch Lernen erworbenes Wissen
Likelihood	Maß, dass die Wahrscheinlichkeit der Ursache beschreibt
Motorik	Vom Gehirn koordinierten Bewegungsabläufe
Myopie	Kurzsichtigkeit
Prävalenz	Kennzahl über die Häufigkeit des Auftretens einer Krankheit

Presbyakusis	Schwerhörigkeit im Alter
Probability of Evidence	Wahrscheinlichkeit, dass ein beobachtetes Ereignis eintritt
Refraktionsfehler	Fehler in der Wahrnehmung, die aufgrund von Brechungsfehlern in der Linse entstehen
Retina	Innere Oberfläche des Auges
Retinopathie	Erkrankung der Netzhaut des Auges aufgrund einer Schädigung kleiner Blutgefäße
Sample Space	Menge von Ausprägungen eines Knotens
View	Ansicht auf einem Bildschirm
Visus	Sehschärfe des Auges
Zerebellär	Das Kleinhirn betreffend

EINFÜHRUNG

Wir leben in einer alternden Gesellschaft. In Europa soll die Anzahl der Personen über 65 Jahren von mehr als 84 Millionen im Jahr 2008 auf etwa 151 Millionen im Jahr 2060 steigen [62]. Weltweit wird erwartet, dass sich die Anzahl der Personen in dieser Altersgruppe, von 524 Millionen in 2011 auf 1,5 Milliarden Menschen in 2050, nahezu verdreifacht [228]. Gleichzeitig erhöht sich auch der Anteil der hochaltrigen Personen, das heißt der Anteil der Personen über 80 Jahre, an der Gesellschaft. Waren im Jahr 2008 etwa 5 % der Personen in der Gesellschaft über 80 Jahre alt, so wird der prognostizierte Anteil bis zum Jahr 2060 auf etwa 14 % steigen [106].

Angesicht dieser Entwicklung wird deutlich, dass Entwickler, Ingenieure und Designer die Anforderungen und Bedürfnisse von älteren Menschen beachten müssen, wenn sie bedienbare Benutzerschnittstellen bzw. Produkte entwerfen wollen, die für alle Teile der Gesellschaft nutzbar sind. Um die Benutzbarkeit der entwickelten Systeme sicher zu stellen, muss die Gebrauchstauglichkeit (engl. Usability) der Systeme getestet bzw. bewertet werden und es müssen aus den Testergebnissen entsprechende Schritte zur Verbesserung des Systems abgeleitet werden können.

1.1 WAS IST USABILITY?

Nach der ISO 9241 Teil 11 [49, S. 4] ist Gebrauchstauglichkeit (engl. Usability) definiert als:

„Das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und zufriedenstellend zu erreichen.“

Definition:
Gebrauchstauglichkeit

Die Gebrauchstauglichkeit bestimmt demnach, wie gut sich ein System durch einen Benutzer bedienen lässt. Um die optimale Gebrauchstauglichkeit sicher zu stellen müssen verschiedene Usability-Engineering-Methoden angewendet werden [163, S. 58 ff]. Die Reihenfolge der Methoden ist im Usability-Engineering Lifecycle (siehe Abbildung 1) dargestellt.

Begonnen wird mit dem Usability-Engineering schon vor dem ersten Systementwurf, indem die Erwartungen der Nutzer, das Nutzungsumfeld und das Verhalten der Nutzer analysiert werden. Die *Analysephase* hat das Ziel, die Anforderungen an das System zu definieren und entsprechende Metriken, inklusive der geplanten Zielwerte, festzulegen,

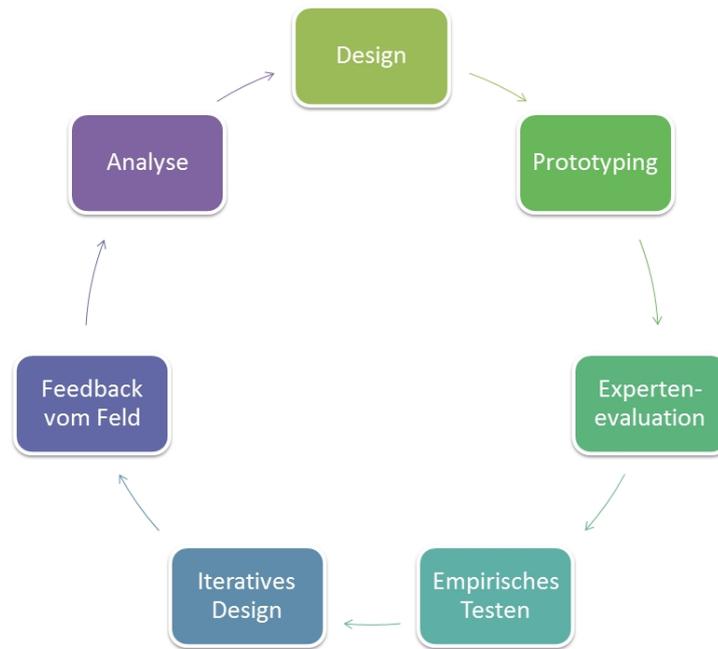


Abbildung 1: Usability-Engineering Lifecycle nach [163]

damit der Erfolg der Umsetzung in späteren Phasen überprüft werden kann.

Der Analysephase folgt die Designphase. Das Ziel der Designphase ist es, viele verschiedene Systementwürfe zu erzeugen, um darauf aufbauend ein Design zu entwerfen, welches die besten Aspekte der verschiedenen Lösungen integriert und im Anschluss an diese Phase weiterentwickelt werden kann. Ist das gewünschte Design definiert folgt die Umsetzung in einen Prototyp.

Prototypen sind zur Überprüfung von Usability-Aspekten notwendig und können für die *Experten-Evaluationen* wie auch für *empirisches Testen* verwendet werden. Unter dem Begriff *Experten-Evaluation* sind Methoden zur Usability-Evaluation zusammengefasst, welche ohne die Einbindung von realen Nutzern auskommen (eine Ausnahme bildet hier der *Pluralistic Usability Walkthrough*). Experten-Evaluationen haben den Vorteil, dass sie bereits in frühen Phasen der Entwicklung eingesetzt werden können.

Unter empirischem Testen sind Versuche mit realen Benutzern zu verstehen. Die Versuche mit realen Nutzern können anhand ihrer Zielsetzung in summative und formative Nutzertests unterteilt werden. Der summative Nutzertest dient dazu die Qualität eines Systems zu bestimmen, während der formative Nutzertest dazu dient Fehler im Design zu finden.

Die Phase des iterativen Designs ist, im eigentlichen Sinn des Wortes, keine eigene Phase, vielmehr ist hierunter das schrittweise Wiederholen der Phasen *Design*, *Prototyping*, *Experten-Evaluation* und *empirisches Testen* zu verstehen. Das Ziel des iterativen Designs ist es, die bestehen-

de Lösung beständig zu verbessern und die Änderungen am Design hinsichtlich möglicher negativer Auswirkungen zu bewerten.

Die letzte Phase im Usability-Engineering Lifecycle ist die Phase der Überführung in den Markt und die Auswertung des Feedbacks vom Feld. Wenn ein System von einer Vielzahl von Nutzern verwendet wird, können viele Informationen gesammelt werden, die für die Verbesserung des Systems relevant sind bzw. die die Grundlage für nachfolgende Systeme legen können.

Um die Usability eines Systems zu bewerten bzw. um Fehler im System zu finden, die die Gebrauchstauglichkeit negativ beeinflussen, können verschiedene Methoden angewendet werden.

1.2 METHODEN ZUR EVALUATION DER USABILITY

Eine erste Unterscheidung für die Auswahl der Methode ist mit dem Ziel der Auswertung verknüpft. Nach Gediga und Hamborg [74] kann eine Evaluierung mit dem Ziel des *Vergleichs zweier Alternativen*, des *Bestimmens der Qualität eines Aspektes* und des *Findens von Fehlern* durchgeführt werden. Während die ersten beiden Zielstellungen eine Methode verlangen, die für die *summative Evaluation* geeignet ist, wird für die dritte Zielstellung eine Methode verlangt, die der *formativen Evaluation* zugeordnet ist. Ist das Ziel der Usability-Evaluation festgelegt, so kann entschieden werden, welche Methoden genutzt werden können um die Ziele der Evaluation zu erreichen.

Nach Sarodnick und Brau [202] können die Usability-Methoden in die Klassen *Gestaltungsrichtlinien*, *formal-analytische Verfahren*, *Inspektionsmethoden* und *Usability-Tests* untergliedert werden. Zusätzlich zu diesen Klassen kann noch die Klasse der Methoden zur automatischen Usability Evaluierung (vgl. Abschnitt 1.4) unterschieden werden.

Die Klasse der *Gestaltungsrichtlinien* lässt sich in *Gestaltungsregeln* (detaillierte Anweisungen), *ergonomische Algorithmen* (Prozeduren die Gestaltungsanforderungen zusammenfassen), *Styleguides* (Dokumente zur Vereinheitlichung von Benutzerschnittstellen), *Standards* (z. B. ISO 9241-11) und *Richtlinien-Sammlungen* (gebündelte Gestaltungsregeln) untergliedern. Die genannten Methoden beschreiben meist Lösungen, wie sie idealerweise umgesetzt werden sollen, können allerdings auch Beispiele geben, wie ein Problem nicht gelöst werden soll (negative Beispiele) [215]. Die Methoden aus der Klasse der *Gestaltungsrichtlinien* können dazu genutzt werden ein System *summativ* wie auch *formativ* zu evaluieren, und sie zählen zu den analytischen Methoden – das heißt, die Methoden funktionieren ohne die Einbindung von Versuchspersonen.

Unter dem Begriff *formal-analytisches Verfahren* sind, nach Sarodnick und Brau [202], *aufgabenanalytische Verfahren* und *Expertenleitfäden* zu verstehen. Ein aufgabenanalytisches Verfahren zerlegen die mit einem System zu lösenden Aufgaben in Teilaufgaben, bis hin zu einzelnen

Handlungsschritten. Eines der bekanntesten aufgabenanalytischen Verfahren ist die GOMS (Goals, Operators, Methods and Selection rules) Methode, die es ermöglicht, Ausführungszeiten für die Lösung einer Aufgabenstellung berechnet zu können, ohne dass das System als Prototyp vorliegt (weiter siehe Abschnitt 1.4). In die Kategorie der Expertenleitfäden gehören alle leitfadenorientierte Prüfverfahren, die die Evaluation anhand einer Checkliste ermöglichen.

In die dritte Klasse, zu den *Inspektionsmethoden*, gehören die Methoden der *heuristischen Evaluationen* und die verschiedenen *Usability-Walkthrough* Varianten. Heuristische Evaluationsverfahren beschreiben die Evaluation der Benutzerschnittstelle anhand von bekannten Usability-Prinzipien (Heuristiken). Nach Möller [163] können Heuristiken hinsichtlich ihrer Detailtiefe in *allgemeine*, *kategorie-spezifische* und *produkt-spezifische* Heuristiken unterschieden werden. Beim Usability-Walkthrough kann zwischen dem *Cognitive Walkthrough* (CTW), dem *Pluralistic Usability Walkthrough* (PUW) und dem *socio-technical Walkthrough* (STWT) unterschieden werden. Der CTW ist die bekannteste Form und das Ziel der Methode ist die Optimierung des Systems dahingehend, dass es schnell von einem unerfahrenen Nutzer erlernt werden kann [154]. Beim CTW werden geplante Handlungsabläufe von Experten durchlaufen und in jedem Schritt wird bewertet a) ob der Nutzer versuchen wird den gewünschten Effekt zu erzielen, b) er erkennen wird dass die korrekte Handlung ausgeführt werden kann, c) ob er erkennen kann, dass die korrekte Handlung den gewünschten Effekt hatte, und d) ob der Nutzer den Fortschritt erkennen kann, wenn er die korrekte Handlung ausgeführt hat. Der PUW beteiligt, anders als der CTW, nicht nur Usability-Experten an der Durchführung, sondern erweitert die Gruppe der Teilnehmer durch repräsentative Nutzer und durch Entwickler [163]. Durch die Einbindung dieser zusätzlichen Evaluatoren erhöht sich die Effizienz der Methode [202]. Der STWT evaluiert nicht nur den Handlungsablauf, sondern das Zusammenspiel zwischen Aufgabe, Nutzer und System, allerdings werden beim STWT keine Fehler in den Benutzerschnittstellen überprüft.

Die von Sarodnick und Brau [202] beschriebene Klasse der *Usability-Tests* unterscheidet sich insofern von den drei vorhergehenden Klassen, als dass es sich hierbei um die erste Methode aus der Gruppe der *empirischen Tests* zur Usability-Evaluation handelt – die vorhergehenden Methoden zählten zu den Experten-Evaluationen. Weiterhin ist der Usability-Test wohl die bekannteste von den bisher aufgeführten Methoden. Beim Usability-Test verwenden Versuchspersonen das System oder einen Prototypen, um reale bzw. realistische Aufgaben zu lösen. Während der Interaktion mit dem System werden die Versuchspersonen beobachtet, und ihr Interaktionsverhalten wird durch sogenannte Logfiles aufgezeichnet. Ein Usability-Test kann *induktiv* oder *deduktiv* sein. Induktive Usability-Tests werden zur formativen

Analyse von Systemen/Prototypen genutzt und dienen dazu Fehler im Systemdesign zu finden. Deduktive Usability-Tests dienen, im Gegensatz zum induktiven Test, zur summativen Evaluation – der sogenannten Ermittlung der Gesamtqualität eines Systems. Da an einem Usability-Test immer Versuchspersonen beteiligt sind, sind die Kosten, die benötigte Zeit und der Aufwand der Durchführung eines solchen Tests hoch [163]. Dies gilt insbesondere, wenn ältere Versuchspersonen am Usability-Test teilnehmen sollen [52] (siehe Abschnitt 1.3.4).

1.3 ÄLTERE VERSUCHSPERSONEN

1.3.1 Begriffsdefinition

Die Begriffe *ältere Versuchspersonen* bzw. *ältere Erwachsene* lassen sich nicht eindeutig definieren [67]. In der Literatur finden sich verschiedene Angaben, ab wann eine Person als *älterer Erwachsener* klassifiziert wird. Einige Studien definieren Personen ab 50 Jahren bereits als *ältere Erwachsene* (z. B. [35, 52, 77, 80, 155]). Demgegenüber stehen anderen Autoren, wie z. B. Neugarten [166, S. 73], der Personen zwischen 55 und 85 Jahren noch als *jüngere ältere Erwachsene* definiert und erst ab über 85 Jahren von den *alten älteren Erwachsenen* spricht.

Da die Definition von *Alter* anhand des chronologischen Alters schwierig ist, werden in der Literatur auch verschiedene alternative Methoden genutzt um den Begriff zu klassifizieren. Dies ist möglich, da *Alter* kein eindimensionales Konstrukt ist, sondern verschiedenen Bedeutungen umfasst, wie das *biologische Alter*, das *psychologische Alter* und das *soziale Alter* [67].

Das *biologische Alter* bezeichnet die Entwicklungsstadien eines Organismus von der Geburt bis zum Tod und ist durch viele Faktoren bestimmt (z. B. Ernährung, Bewegung und Alkoholkonsum) [207, S. 398], wohingegen der Begriff des *psychologischen Alters* Faktoren wie das subjektiv Empfinden der eigenen Gesundheit und die kognitive Leistungsfähigkeit eines Menschen [241, S. 428] umfasst. Das *soziale Alter* definiert sich über die Zugehörigkeit zu einer durch die Gesellschaft abgegrenzten Altersphasen bzw. Altersgruppen [207, S. 398]. Alle genannten Begriffe eint der Umstand, dass sie mit dem chronologischen Alter in Beziehung gesetzt werden. Dies geschieht, weil das chronologische Alter in vielen Gesellschaften als Vertreter (im Sinne von Proxy) für die biologische Reife, die psychologische Entwicklung und die Zugehörigkeit zu einer sozialen Kategorie oder Lebensphase steht [213].

Für diese Arbeit wird der Begriff der *älteren Versuchspersonen* bzw. *ältere Erwachsene* am chronologischen Alter orientiert und relativ weit gefasst, das bedeutet, dass schon Personen ab 50 Jahren zu dieser Gruppe gezählt werden. Diese Entscheidung resultiert aus dem Umstand, dass verschiedene altersbedingte Einschränkungen bereits mit

Definition:
Ältere
Versuchspersonen

Mitte 40 Jahren beginnen [67, S. 61] und ab 50 diese Einschränkungen in den entsprechenden Altersgruppen stark zunehmen (z. B. die Verbreitung der Trübung der Augenlinse – der sogenannte Katarakt [157]).

1.3.2 Einschränkungen

Eine große Herausforderung für Usability-Tests unter Beteiligung älterer Versuchspersonen ist die erhebliche Heterogenität bezüglich der sensorischen, kognitiven und motorischen Fähigkeiten, welche mit höherem chronologischem Alter noch zunimmt [82], und die aus dieser Heterogenität resultierenden unterschiedlichen Interaktionsprobleme [67].

Für die Interaktion mit einem System zählen die *visuelle Wahrnehmung* und die *auditive Wahrnehmung* zu den wichtigsten altersbedingten Einschränkungen in der sensorischen Perzeption [67]. Nach Holzinger [103] sind die nachfolgend genannten visuellen wie auch auditiven Eigenschaften a) besonders wichtig in der Interaktion mit einem System und b) von altersbedingten Einschränkungen betroffen. Für die visuelle Wahrnehmung sind dies:

1. die Sehschärfe (die Fähigkeit feine Details wahrzunehmen),
2. die Nahakkommodation (die Fähigkeit die Brechkraft der Linse anzupassen um nahe Objekte wahrnehmen zu können),
3. die Farbwahrnehmung (die Fähigkeit verschiedene Farben zu unterscheiden),
4. die Kontrastwahrnehmung (die Fähigkeit Unterschiede in der Helligkeit wahrzunehmen),
5. die Dunkeladaptation (die Fähigkeit sich schnell an dunkle Umgebungen anzupassen),
6. die Blendempfindlichkeit (die Fähigkeit des Auges sich schnell an helles Licht anzupassen),
7. die Lichtempfindlichkeit (das Level des minimal benötigte Lichts),
8. die Wahrnehmung von Bewegung und
9. das periphere Sehen (das Sehen außerhalb des Bereichs des schärfsten Sehens).

Im Bereich der auditiven Eigenschaften sind nach [103] die *Hörschwelle*, die *auditorischer Lokalisierung* und die *Fähigkeit Sprache in lauten Umgebungen wahrzunehmen*, von entscheidender Bedeutung.

Neben den sensorischen altersbedingten Einschränkungen existieren auch kognitive altersbedingte Einschränkungen. Der Bereich der

kognitiven Beeinträchtigungen umfasst die *Informationsverarbeitung*, das *Gedächtnis*, die *Reaktionsfähigkeit* und die *Koordination* [15, S. 161].

In verschiedenen Versuchen wurde beobachtet, dass ältere Personen weniger Informationen zeitgleich verarbeiten können, und dass die Verarbeitung länger dauert [204, S. 70]. Ebenso konnte beobachtet werden, dass die Speicherung neuer Informationen im Gedächtnis langsamer erfolgt, was Lernprozesse verlängert [150, S. 357]. In der gleichen Weise kommt es durch die verlangsamte Reaktionszeit dazu, dass es länger dauert bis ältere Menschen auf einen Reiz reagieren [204, S. 70]. Durch die Einschränkungen im Kurzzeitgedächtnis kommt es weiterhin zu Problemen in der Koordination von verschiedenen Aufgaben, so dass parallele Aufgaben langsamer ausgeführt werden [104, S. 152].

Die *kognitive Leistungsfähigkeit* (Intelligenz) lässt sich in die sogenannte *fluide* und *kristalline Intelligenz* untergliedern [15, S. 176]. Die kristalline Intelligenz beschreibt die über alle Lebensphasen erworbenen Fähigkeiten und das gespeicherte Wissen. Im Gegensatz zur fluiden Intelligenz unterliegt die kristalline Intelligenz kaum einem altersbedingten Abbau [150, S. 350]. Die fluide Intelligenz beschreibt die Fähigkeit des Gehirns neue Informationen zu erlernen bzw. wahrgenommene Informationen zu verarbeiten und ist einem altersbedingten Verfall unterworfen [137, S. 359], was z. B. zu Problemen beim Erlernen neuer Interaktionsformen führt [86].

Die motorischen Einschränkungen lassen sich einerseits in *funktionelle* und *strukturelle Veränderungen* und andererseits in die Kategorien *Beweglichkeit*, *Kraft* und *Fingerfertigkeit* unterteilen [15]. Zur Klasse der funktionellen Einschränkungen in der Beweglichkeit gehören z. B. Veränderungen in den Gelenken, so dass diese an Festigkeit und Beweglichkeit verlieren [204, S. 67]. Rheumatische Erkrankungen wie Arthrose oder rheumatische Arthritis gehören hingegen zu den strukturellen Einschränkungen in der Bewegung [15, S. 172]. Hinsichtlich der Kraft kann festgestellt werden, dass die Muskelkräfte im Alter abnehmen und dies zu Schwierigkeiten in der Interaktion mit einem System führen kann [15, S. 173]. Durch funktionelle und strukturelle Veränderungen in der Fingerfertigkeit wird der Tastsinn eingeschränkt, was zu Fehlern in der Bedienung führt, besonders in dem Fall, wenn die Bedienelemente klein oder dicht beieinander angeordnet sind [15, S. 175].

Neben den altersbedingten Einschränkungen sind auch fehlende Erfahrungen im Umgang mit modernen Eingabegeräten eine Quelle für Fehler während der Interaktion [87]. Die fehlenden Erfahrungen im Umgang mit modernen Eingabegeräten lassen sich dadurch begründen, dass ältere Menschen meist Technik nutzen, die von ihren Kindern abgelegt wurde; die älteren Menschen sind somit meist eine Technikgeneration hinter dem aktuellen Stand der Technik zurück [86]. Problematisch hierbei ist, dass für ältere Menschen das Erlernen

von neuen Interaktionsformen eine große Hürde darstellt, weil das Erlernen neuer Fähigkeiten hohe Anforderungen an die fluide Intelligenz stellt [86], und wie bereits dargestellt wurde, ist fluide Intelligenz stark von altersbedingten Einschränkungen betroffen.

Ein Ansatz zur Erklärung von Interaktionsproblemen – die nicht mit altersbedingten Einschränkungen erklärt werden können – ist der Ansatz der Technikgenerationen von Sackmann und Weymann [198]. Das Konzept der Technikgeneration geht davon aus, dass die Personen der entsprechenden Personengruppe durch die Geräte geprägt werden, die sie in jüngeren Jahren (bis zum 25. Lebensjahr) verwenden. Sackmann und Weymann nehmen an, dass die in dieser Phase erworbenen Erfahrungen im Umgang mit Technik auch später noch Einfluss auf das Interaktionsverhalten haben. Die Problematik der frühen Prägung zeigt sich besonders, wenn neue Technologien auf den Markt kommen, die ältere Technologien ablösen, und wenn die in früherer Zeit erworbenen *mentalen Modelle* (siehe den nachfolgenden Abschnitt) von der Funktionsweise eines Systems bzw. eines Gerätes nicht mehr passen, um die neue Technologie bedienen zu können [259].

1.3.3 Mentale Modelle

Was ist ein *mentales Modell*? Die Definition von mentalen Modellen in dieser Arbeit stützt sich auf die Ausführungen von Carroll und Olson (1987). Nach Carroll und Olson ist ein mentales Modell wie nachfolgend beschrieben definiert:

Defintion:
Mentale Modelle

„[The] user’s mental model of a system is [...] defined as a rich and elaborate structure, reflecting the user’s understanding of what the system contains, how it works, and why it works that way.“[121, S. 12]

Ein mentales Modell ist demzufolge etwas, was Wissen voraussetzt und von Nutzer zu Nutzer unterschiedlich sein kann. Nach Norman [76, S. 8] haben mentale Modelle die nachfolgend aufgeführten Eigenschaften:

1. Mentale Modelle sind nicht vollständig.
2. Die Möglichkeit von Menschen ihre mentalen Modelle auszuführen ist begrenzt.
3. Mentale Modelle sind nicht stabil.
4. Mentale Modelle haben keine festen Grenzen und verschwimmen zwischen ähnlichen Systemen.
5. Mentale Modelle sind nicht rational. Manche Personen führen unnötige Operationen aus, obwohl sie wissen, dass diese unnötig

sind. Sie tun dies, weil der Aufwand für die Ausführung der physischen Bewegung geringer ist als der mentale Aufwand, wenn sie das Verhalten ändern würden.

6. Mentale Modelle sind sparsam. Statt komplizierter, systemspezifischer Operationen, welche den physischen Aufwand reduzieren, werden einfache, allgemeine Operationen im mentalen Modell gespeichert, die für viele verschiedene Systeme anwendbar sind.

Ein mentales Modell kann Software (das System) und Hardware (das Eingabegerät) betreffen. Zur besseren Unterscheidung werden nachfolgend die Begriffe *System Mental Model* (SMM) und *Device Mental Model* (DMM) eingeführt. Das SMM beschreibt die Vorstellung des Nutzers vom System, demzufolge wird das SMM definiert als:

Die Annahme des Nutzers bezüglich der Reaktion des Systems auf eine Eingabeaktion.

Definition:
System Mental Model

Das DMM beschreibt die Vorstellung des Nutzers vom Eingabegerät, folglich kann das DMM definiert werden als:

Die Mutmaßung des Nutzers wie das Eingabegerät bedient werden muss, um eine Eingabeaktion auszulösen.

Definition:
Device Mental Model

Wird der Ansatz der Technikgenerationen mit der Definition und den Eigenschaften von mentalen Modellen und den altersbedingten Einschränkungen verbunden, so zeigen sich interessante Aspekte, die die Interaktion mit einem System und somit die Usability negativ beeinflussen können.

In [87] konnte Hanson zeigen, dass fehlende Erfahrungen im Umgang mit modernen Eingabegeräten, bedingt durch Fehler im DMM, zu Fehlern in der Interaktion führen. Weiterhin konnte Hanson zeigen, dass durch kognitive altersbedingte Einschränkungen die Fähigkeit neue Eingabeformen zu erlernen, neue DMM zu bilden oder bestehen DMMs anzupassen, stark beeinträchtigt wird. Durch die Beeinträchtigung des Lernprozesses können die mentalen Modelle nicht oder nur langsam an die neuen Eingabegeräte angepasst werden. In dieser Zeit der Anpassung werden häufig Interaktionsmuster auf ein Eingabegerät angewendet, die von anderen ähnlichen Eingabegeräten übernommen bzw. abgeleitet wurden (siehe Punkte 4. in der Aufzählung oben). Sind diese übernommenen Eingabeformen nicht passend für das neue Eingabegerät, so kann es zu Bedienfehlern kommen, wenn das Interaktionsmuster nicht verarbeiten kann [68, 178].

1.3.4 Usability-Tests mit älteren Versuchspersonen

Traditionelle Usability-Tests unter Beteiligung von älteren Versuchspersonen sind schwierig umzusetzen und können hohe Kosten verursachen [52]. Nach Fisk et al. [67, S. 171 ff] ist es für einen Usability-Test

von entscheidender Bedeutung, dass Nutzer eingeladen werden, die repräsentativ für die geplanten Zielgruppen stehen. Deshalb sollen die Versuchsteilnehmer nicht nur nach dem chronologischen Alter ausgewählt werden, sondern hinsichtlich ihrer Fähigkeiten, Vorerfahrungen, Gesundheitszustand, Leistungsfähigkeit und Einschränkungen. Nach Hawthorn [92] ist die Erfüllung der genannten Voraussetzungen allerdings sehr schwierig, da an den entsprechenden Tests meist nur die physisch und kognitiv fitten älteren Personen teilnehmen, was zu falschen Ableitungen aus den Testergebnissen führen kann.

Stellen die höheren Aufwände und die Auswahl an geeigneten Versuchspersonen keine Probleme dar, so können durch die Wahl der Evaluierungsmethode kritische Probleme entstehen. Nachfolgende werden drei empirische Testmethoden (*Thinking aloud*, *Nutzertagebücher* und *Fokusgruppen*) vorgestellt, bei denen die Beteiligung von älteren Versuchspersonen zu Herausforderungen führen kann.

Die erste genannte Methode ist *Thinking aloud*, die Methode des *Lauten Denkens*. Bei dieser Methode muss die Versuchsperson, während sie mit dem System interagiert, ihre Gedanken verbalisieren (aussprechen) [202]. Durch *Thinking aloud* können mit wenigen Versuchspersonen viele hilfreiche Informationen gesammelt werden. Die wesentlichen Nachteile der Methode sind die höhere mentale Belastung, durch das gleichzeitige Sprechen und Handeln, und der Umstand, dass die Versuchspersonen durch das Bewusstmachen der eigenen Handlung einige Fehler vermeiden. Für ältere Versuchsteilnehmer ist diese Methode oftmals nicht geeignet. So erläutern Fisk et al. [67], dass *Thinking aloud* durch die höhere mentale Belastung für ältere Menschen ungeeignet ist, besonders wenn ältere Versuchspersonen mit kognitiven Einschränkungen für sie unbekannte Benutzerschnittstellen evaluieren sollen. Auch *retrospektives Thinking aloud* (realisiert durch Videoaufzeichnungen) ist nicht geeignet, da viele ältere Versuchspersonen sich häufig nicht daran erinnern können, welche Gedanken sie während der Lösung der Aufgabe hatten bzw. weshalb Probleme während der Interaktion auftraten [52].

Ähnliche Schwierigkeiten können beobachtet werden, wenn ältere Personen im Rahmen einer Studie *Nutzertagebücher* führen sollen, um ihr tägliches Erleben im Umgang mit dem System zu dokumentieren. So zeigte sich in der Untersuchung von Czerwinski et al. [46], dass das Führen des Tagesbuches den natürlichen Tagesablauf erheblich störte und dass speziell ältere Versuchspersonen keine Zeit hatten das Tagebuch auszufüllen, wenn sie Probleme in der Interaktion mit einem System hatten. Ein weiteres Problem der Methode resultiert aus motorischen Einschränkungen von älteren Nutzern. So wurde von Dickinson et al. [52] beobachtet, dass Nutzer mit motorischen Einschränkungen vermeiden viel zu schreiben, da das Aufschreiben als sehr anstrengend empfunden wird.

Werden *Fokusgruppen* mit älteren Teilnehmern durchgeführt, so können durch altersbedingte Einschränkungen Probleme entstehen, die die Durchführung der Methode erschweren [11]. Bei der Methode Fokusgruppe werden 6 bis 9 Teilnehmer eingeladen und diese werden dazu aufgefordert, über verschiedene Systemlösungen zu diskutieren [163]. In Fokusgruppen können altersbedingte Hörprobleme oder auch Verständnisprobleme zu Kommunikationsproblemen führen, die eine effiziente Zusammenarbeit verhindern [168]. Altersbedingte Einschränkungen in der Aufmerksamkeit können dazu führen, dass es den älteren Teilnehmer schwerer fällt, auf das Thema fokussiert zu bleiben [12]. Mit dem Alter assoziierte kognitive Einschränkungen, wie die kognitive Verlangsamung (general slowing theory; [199]), können dazu führen, dass es zu Brüchen im Themenverlauf kommt, da ältere Versuchspersonen oftmals lange über Fragen bzw. Themen nachdenken und nach Abschluss ihrer Überlegungen verspätet antworten, selbst wenn in der Zwischenzeit das Thema gewechselt wurde [11].

Wie dargestellt ist es im Allgemeinen nicht einfach, die Usability von Systemen mit älteren Versuchspersonen zu testen. Durch die Heterogenität und altersbedingten Einschränkungen der älteren Versuchspersonen sind empirische Usability-Tests teuer und aufwändig [52]. Weiterhin sind die erzielten Ergebnisse aufgrund der gewählten Methode und den teilnehmenden Versuchspersonen nicht immer repräsentativ für die gewünschte Zielgruppe [92]. Es wird somit eine Methode benötigt, die hilft die Usability für so problematische Gruppen zu evaluieren, ohne dass repräsentative Versuchspersonen eingeladen werden müssen.

Die Methode der Automatischen Usability-Evaluation (AUE) ist ein vielversprechender Ansatz, um die Gebrauchstauglichkeit von Benutzerschnittstellen für die Gruppe der älteren Menschen zu evaluieren bzw. zu optimieren [119] – so können z. B. Simulatoren dazu genutzt werden, den Designern einen Eindruck davon zu vermitteln, wie sich eine Einschränkung auswirken kann [81].

1.4 AUTOMATISCHE EVALUIERUNG DER USABILITY

Was bedeutet *Automatische Evaluierung der Usability*? Nach Shewanown [214] handelt es sich bei der automatischen Evaluierung der Usability um eine Methode, die die traditionellen Verfahren ergänzt und eingesetzt werden kann, wenn *Menschen, Zeit* oder *finanzielle Ressourcen* knapp sind.

Die automatische Evaluierung der Usability ersetzt somit nicht die klassischen Methoden, sondern generiert einen Mehrwert, der hilft Zeit und Geld zu sparen. Der Mehrwert der Methode resultiert aus dem Einsatz in frühen Phasen des Usability-Engineering Lifecycles (siehe Abbildung 2). Mit Hilfe der automatischen Evaluierung der Usability können Usability-Fehler im Design frühzeitig gefunden und

Definition:
*Automatische
Evaluierung der
Usability*

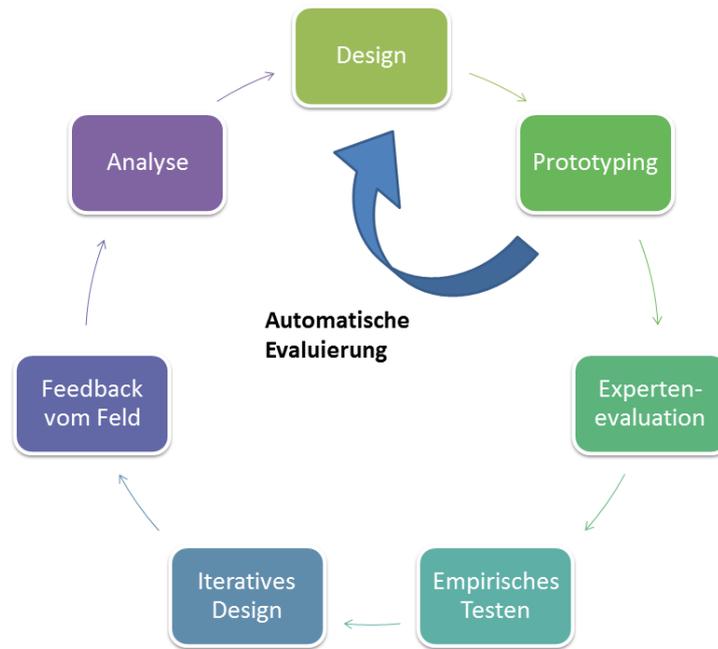


Abbildung 2: Usability-Engineering Lifecycle mit automatischer Evaluierung

behooben werden. Nach Gilb und Finzi [79] kann durch die frühzeitige Behebung der Fehler viel Geld gespart werden, da gilt, je früher ein Usability-Fehler gefunden und behoben wird, desto günstiger ist die Behebung des Fehlers.

In Bezug auf Usability-Tests mit älteren Nutzern bedeutet dies, dass der Entwickler weniger ältere Versuchspersonen einladen muss, um das Design formativ zu testen. Im Idealfall können durch die automatische Evaluierung alle Usability-Fehler gefunden und vor dem empirischen Test behoben werden. Der abschließende empirische Test dient für diesen Fall nur noch zur summativen Evaluation der Usability der Benutzerschnittstelle. Da für die automatische Evaluation verschiedene Methode und Werkzeuge existieren, werden nachfolgende die wichtigsten Vertreter in der chronologischen Abfolge ihrer Veröffentlichung vorgestellt. Für jede Methode wird ausgeführt wie diese funktioniert und wann sie eingesetzt werden kann.

1.4.1 Methoden und Werkzeuge

Die Evaluation von Mensch-Maschine Schnittstellen ohne die Einbeziehung von realen Nutzern ist keine neue Idee. Schon im Jahr 1974 wurde von Bonney et al. [20] das System SAMMIE (System for Aiding Man Machine Interaction Evaluation) vorgestellt. SAMMIE ist eines der ersten Systeme, mit denen die Mensch-Maschine Interaktion in sehr frühen Phasen der Entwicklung evaluiert werden kann. Das Werkzeug wurde für die Automobilindustrie entwickelt, um am Computer

die Modelle neuer Bedienelemente im Auto (Schalter, Knöpfe, usw.) zu testen. Die Version von 1974 enthielt bereits ein (Nutzer-)Modell eines durchschnittlichen Arbeiters, ein Werkzeug um den Arbeitsplatz virtuell zu modellieren, eine eigene Sprache um die Aufgabenstellung für das Nutzermodell zu definieren, heuristische Regeln um die Anordnung der Bedienelemente zu überprüfen, verschiedene Algorithmen um die Effektivität zu berechnen, und ein System um die Interaktion zwischen Mensch und System zu simulieren. Simulation heißt in diesem Zusammenhang, dass das Werkzeug berechnete, ob die zur Lösung der Aufgabe benutzten Bedienelemente erreichbar und ob alle notwendigen Informationen von der Position des Nutzermodells aus sichtbar waren. Von Kingsley et al. [132] wurde im Jahr 1981 die Anwendbarkeit von SAMMIE auf den Bereich der Mensch-Computer Interaktion gezeigt, allerdings bezog sich die Evaluation nicht auf das verwendete Programm (Software) sondern auf die Anordnung der Eingabegeräte (Hardware) auf dem Schreibtisch. Von Porter et al. [180] wurde im Jahr 2004 eine Erweiterung der Software vorgestellt, die es ermöglicht die Bedienbarkeit eines Systems durch Menschen mit Einschränkungen zu simulieren.

Die erste bekanntere Methode im Bereich der Mensch-Computer-Interaktion wurde im Jahr 1980 von Card und Kollegen [31] vorgestellt; es handelt sich hierbei um das Keystroke-Level Model (KLM). Das KLM ist eine vereinfachte (frühe) Version der GOMS (Goals, Operators, Methods and Selection rules) Methode, welche ebenfalls von Card et al. [30] entwickelt wurde. KLM und GOMS sind dazu geeignet die Ausführungszeiten zu berechnen, die nötig sind, um eine Aufgabe mit dem gegebenen System zu lösen. Beide Methoden machen die Annahme, dass die Aufgabe von einem nicht eingeschränkten Experten bearbeitet wird und dass dieser die ihm gestellte Aufgabe fehlerfrei lösen kann. Zur Berechnung der Ausführungszeiten werden bei KLM und auch bei GOMS die *Ausführungsschritte* und andere elementare Ereignisse (z. B. *mentale Vorbereitung* und *Systemreaktionszeiten*) gezählt, mit vorgegebenen Zeiten verknüpft, und die so definierten Zeiten summiert. Beide Methoden eignen sich insbesondere dazu, die Ausführungszeiten von verschiedenen Systemen bzw. alternativen Lösungen miteinander zu vergleichen. Für GOMS und KLM existieren Modellvarianten, die eine Berechnung von Ausführungszeiten älterer Experten erlauben ([67, S. 217 ff] bzw. [117]).

ACT-R (Adaptive Control of Thought-Rational [3]) ist eine rechnergestützte kognitive Architektur, welche dazu genutzt werden kann die menschliche Kognition zu simulieren. Das Ziel von ACT-R ist es das Zusammenspiel von menschlichen kognitiven Prozessen zu verstehen und nachbilden zu können. Um dieses Ziel zu erreichen nutzt ACT-R verschiedene Module und Puffer, um den Ablauf der

Informationsverarbeitung (die Abfolge der Aktivierungsprozesse) in der ACT-R-Laufzeitumgebung nachzubilden. Mit Hilfe von ACT-R kann man Vorhersagen über das Nutzerverhalten und die Nutzerperformanz, nicht nur von jüngeren Experten, treffen. So wurde von Salvucci et al. [201] modelliert, wie sich altersbedingte kognitive Einschränkungen auf das Fahrverhalten auswirken. Jastrzemski et al. [115] stellten in 2010 verschiedene ACT-R Modelle vor, die dazu geeignet sind, die Performanz von älteren Nutzern bei der Bedienung von Mobiltelefonen vorherzusagen und theoretisch fundierte Erklärungen für die Unterschiede zwischen jüngeren und älteren Nutzern zu liefern. Neben diesen Vorteilen existieren auch ein paar Nachteile. So ist zur Programmierung der ACT-R-Modelle ein methoden-spezifischer Syntax notwendig, welcher vom Entwickler erlernt werden muss. Dieser Umstand schränkt die Anwendbarkeit in der Praxis stark ein. Da ACT-R in der wissenschaftlichen Forschung starke Anwendung findet, wird es aktiv weiterentwickelt. Neben ACT-R existieren noch weitere kognitive Architekturen (z. B. SOAR [143], EPIC [131], CAPS [122] und CORE [239]), welche jedoch wesentlich seltener im Bereich der Mensch-Maschine Interaktion genutzt werden.

Um die Anwendung von ACT-R im Bereich der Mensch-Maschine Interaktion zu vereinfachen wurde von John und Kollegen die Software CogTool (CT) entwickelt [120]. Das CT bietet dem Modellierer eine grafische Oberfläche um einen Prototyp (Modell) des zu testenden Systems zu erstellen. Ist das Modell erstellt, definiert der Modellierer den zu erwartenden Interaktionspfad durch das System. Das CT übersetzt die von Modellierer aufgezeichneten Interaktionsschritte und wandelt diese in mehreren Schritten in ACT-R-Code um. Der so erzeugte Code wird anschließend von der ACT-R-Laufzeitumgebung ausgeführt und die Zeit für die Ausführung der Schritte berechnet. Ähnlich GOMS und KLM berechnet das CT unter Verwendung von ACT-R die Ausführungszeiten für einen Experten. Zwar wurde im Jahr 2010 von John und Jastrzemski [119] gezeigt, dass das CogTool, nach Anpassung der zugrundeliegenden Modelle, in der Lage ist, die Ausführungszeiten von älteren Personen bei der Bedienung von Mobiltelefonen vorherzusagen, allerdings ist dies im Allgemeinen in den Simulationen nicht vorgesehen. Weiterhin ist nicht vorgesehen, dass Interaktionsfehler auftreten oder dass eine Einschränkung den Interaktionsverlauf beeinträchtigt. Zweifellos steht es dem Entwickler offen auch fehlerbehaftete Pfade und deren Ausführungszeiten zu simulieren, allerdings wird das CT hierfür meist nicht genutzt.

Neben dieser originalen Version existiert eine erweiterte Variante des Werkzeugs – der sogenannte CogTool-Explorer (CTE). Der CTE versucht die Ausführungszeiten von Novizen vorherzusagen, indem es Kaskaden von Augenbewegungen simuliert und, unter Verwendung von SNIF-ACT 2.0 [70], einen semantischen Vergleich von Aufgaben-

wissen und Wörtern in der grafischen Oberfläche zur Auswahl des nächsten Interaktionsobjektes durchführt. Durch die automatische Auswahl des nächsten Interaktionsobjektes ist es für den Modellierer nicht notwendig einen Interaktionspfad vorzugeben, was den Vorteil bietet, dass Fehler in der Interaktionslogik gefunden werden können, da die so generierten Interaktionspfade von den vorgesehenen Pfaden abweichen können. Ein Problem bei der Verwendung des CTE ist, dass die Benutzeroberfläche und das Aufgabenwissen in englischer Sprache vorliegen müssen, da der für den Vergleich genutzte Webservice nur englische Wörter verarbeiten kann. Seit Dezember 2013 steht der Quellcode von CT und CTE über Github¹ zur Verfügung.

Ein weiteres Werkzeug, das den Entwickler bei der Evaluierung der Usability eines Systems unterstützt, ist der HUI (Handheld User Interface) Analyzer [9]. Der HUI Analyzer kann dazu genutzt werden auf .Net basierende Windows-Mobile-Anwendungen (semi-) automatisch zu evaluieren bzw. Ergebnisse aus empirischen Versuchen zu visualisieren. Um eine App zu evaluieren muss der Entwickler die fertige App in das Werkzeug laden. Ist die App in das Werkzeug geladen definiert der Entwickler den erwarteten Interaktionspfad um eine Aufgabe zu lösen – ein Vorgehen welches dem CogTool ähnelt. Als nächsten Schritt lädt der Entwickler die in einem empirischen Versuch ermittelten Interaktionspfade in das Werkzeug. Basierend auf diesen beiden Eingaben wird vom HUI Analyzer berechnet wie hoch die Übereinstimmung des erwarteten Interaktionsverlaufs mit den tatsächlichen Interaktionsverläufen ist. Zusätzlich werden Heatmaps über häufig genutzte Interaktionselemente gelegt, um besonders wichtige Interaktionsobjekte hervorzuheben. Können keine empirischen Interaktionspfade bereitgestellt werden, so kann die App gegen, vom Entwickler zu bestimmende, Usability Metriken (Heuristiken) geprüft werden. Die Weiterentwicklung des HUI Analyzers wurde zwischen eingestellt, so dass nicht bekannt ist, ob die Software noch genutzt oder für die Vorhersage für die Nutzergruppe der älteren Personen verwendet werden kann.

Die MeMo-Werkbank ist ein weiteres Werkzeug zur (semi-) automatischen Evaluierung der Usability in frühen Phasen der Entwicklung [60]. Für die Verwendung der Werkbank muss kein fertiges System vorliegen, vielmehr wird in der Werkbank ein Modell bzw. werden verschiedene Modelle eines Systems konstruiert, welche im Anschluss von Nutzermodellen (virtuellen Nutzern) verwendet werden, um definierte Aufgaben mit einem definierten Aufgabenwissen zu lösen. Ähnlich dem CTE werden die Pfade durch das System nicht vom Modellierer vorgegeben, sondern vom Werkzeug anhand des Aufgabenwissens berechnet. Im Unterschied zum CTE berechnet die

¹ <https://github.com/cogtool/cogtool>

MeMo-Werkbank pro Simulation meist mehrere Iterationen und stellt dem Entwickler die zusammengefassten Ergebnisse zur Verfügung. Durch die Vielzahl der Iterationen pro Simulation werden – meist bedingt durch vorhandene Usability-Fehler – viele Abweichungen vom optimalen Interaktionspfad gefunden. Jede Abweichung kann vom Modellierer betrachtet und bewertet werden, um zu entscheiden, ob der Grund für die Abweichung ein Usability-Problem ist, und falls ja, wie schwerwiegend das Problem ist. Da die MeMo-Werkbank die technische Basis für diese Arbeit bildet, wird sie in Kapitel 2 detailliert beschrieben.

Der Biswas-Simulator dient dazu, die Auswirkungen von Beeinträchtigungen auf das Interaktionsverhalten zu simulieren, um den Designer für die Bedürfnisse und Anforderungen von Menschen mit Handicap zu sensibilisieren [17]. Wie beim CT, CTE bzw. bei der MeMo-Werkbank modelliert der Entwickler das zu testende System mit Hilfe des Werkzeugs. Anders als das CTE und die MeMo-Werkbank nutzt der Biswas-Simulator allerdings das *dual-space model* (DSM) von Rieman et al. [188] zur Simulation der Interaktionspfade bzw. zur Vorhersagen der Performanz von Novizen [18]. Das DSM beschreibt zwei kognitive Modelle für exploratives Lernen, wobei *trail and error* die erste und *nach Hilfe fragen* die zweite Strategie darstellt [188]. Die Verwendung des DSM sieht im Fall des Biswas-Simulators konkret so aus, dass das Nutzermodell erst den Modellierer fragt, wie der Pfad zur Lösung der Aufgabe aussieht (ähnlich dem CT). Hier hat der Modellierer die Möglichkeit den Pfad vorzugeben, und somit dem Vorgehen des CTs zu entsprechen, oder keine Vorgaben zu machen. Werden dem Nutzermodelle vom Modellierer keine Vorgaben gemacht, so sucht es im Interface nach Beschreibungen (Labeln), welche dem Aufgabenwissen entsprechen und wählt eine Interaktion entsprechend den Beschreibungen aus [17]. Kann keine passende Beschreibung gefunden werden, so wählt das Nutzermodell eine Interaktion aus, welche es in einen schon bekannten Systemzustand bringt. Mit diesem Vorgehen ist der Biswas-Simulator in der Lage verschiedene Sequenzen von Interaktionen zu simulieren. Zusätzlich zur Simulation der verschiedenen Interaktionspfade kann der Simulator noch die Auswirkungen von Einschränkungen darstellen, indem die verwendeten Screenshots bzw. Grafiken durch Algorithmen so verändert werden, dass der Designern einen Eindruck davon erhält, wie sich die ausgewählten Einschränkungen auf die Wahrnehmung auswirken.

Ähnlich dem Biswas Simulator simuliert auch der Disability Impairment Approximation Simulator (DIAS) die Auswirkungen von körperlichen, sowie kognitiven, Einschränkungen auf die Interaktion [78]. Vom DIAS existieren zwei verschiedene Varianten: eine eigenständige Programmversion und ein NetBeans Plug-In zur Integration

in eben jene Entwicklungsumgebung. Die eigenständige Version kann dazu genutzt werden, bereits vorhandene JAVA-Programme zu laden und die Auswirkungen von visuellen Einschränkungen zu simulieren, so dass der Entwickler einen Eindruck davon erhält, welche Probleme durch die Einschränkungen entstehen. Um den Entwickler während der Entwicklungsphase zu unterstützen kann das Netbeans Plug-In benutzt werden. Beide Varianten bieten für einige wenige ausgewählte Einschränkungen einen Hinweis darauf, wie viele Personen in der Bevölkerung von den entsprechenden Einschränkungen betroffen sind, allerdings sind die Zusammenhänge zwischen einzelnen Einschränkungen nicht klar. Zur aktuellen Verwendbarkeit der Werkzeuge kann festgestellt werden: Das Netbeans Plug-In wurde 2013 letztmalig aktualisiert, die Arbeiten an der eigenständigen Programmversion nach Abschluss des ACCESSIBLE-Projektes² im Jahr 2010 eingestellt, somit kann davon ausgegangen werden, dass die Software nicht mehr dem aktuellen Stand der Technik entspricht.

Zusammenfassend kann festgehalten werden: Im Bereich der AUE existieren verschiedene Lösungen die unterschiedliche Ziele verfolgen. Einige Methoden werden genutzt um Ausführungszeiten zu berechnen, andere um Usability-Probleme zu finden. Um Usability-Probleme zu evaluieren werden Nutzermodelle benötigt, die reale Nutzer repräsentieren und bestimmte Eigenschaften haben. Da der Begriff *Nutzermodell* bisher nicht eingeführt wurde, wird im nachfolgenden Abschnitt erläutert, was unter dem Begriff zu verstehen ist, wie Wissen über Nutzermodelle abgespeichert werden kann, und welche offenen Probleme existieren.

1.4.2 Nutzermodelle

Der Begriff *Nutzermodell* hat im Bereich von HCI zwei verschiedene Bedeutungen [128].

Die erste Bedeutung verweist auf Modelle, die Nutzer beschreiben, die ein System verwenden. Diese erste Variante von Nutzermodellen dient häufig dazu Designern bzw. Systementwickler dabei zu helfen, eine Entscheidung zwischen verschiedenen alternativen Benutzeroberflächen zu treffen. Dies ist möglich, da solche Nutzermodelle die Eigenschaften bzw. Einschränkungen der potenziellen Nutzer widerspiegeln. Die bekannteste Form dieses Typs sind die *Personae*. Unter *Personae* sind anonymisierte Modelle von Personen zu verstehen, die das Verhalten und die Motivationen von realen Personen widerspiegeln, so wie diese in Versuchen beobachtet worden sind [41].

Die zweite Variante beschreibt einen Typ von Modell, welcher in ein System integriert ist. Ein Nutzermodell von diesem Typ wird vom System dazu genutzt, um sich an die verschiedenen Nutzer

² <http://www.accessible-eu.org/>

bzw. an die Eigenschaften und Präferenzen der Nutzer anzupassen bzw. zukünftige Handlungen des Nutzers vorherzusehen. Ein Nutzermodell dient somit als Vorlage zur Adaption des Systems an den Nutzer. Die bekannteste Form für diesen Typ von Nutzermodellen befindet sich in vielen Empfehlungssystemen, wie z. B. dem Amazon-Empfehlungssystem.

Im Rahmen dieser Arbeit ist der Begriff *Nutzermodell* in Anlehnung an die erste Definition von Kass [128] wie folgt festgelegt:

Definition:
Nutzermodell

Ein Nutzermodell beschreibt alle Eigenschaften eines Nutzers, die sich auf die Interaktion mit einem System auswirken können. Hierbei ist zu beachten, dass die Eigenschaften selbst und der Zusammenhang zwischen verschiedenen Eigenschaften realistisch abgebildet werden, und dass eine Sammlung von Nutzermodellen die Heterogenität von Nutzern in der gesamten Gesellschaft oder in einem Teil der Gesellschaft repräsentativ reproduzieren kann.

Werden die in Abschnitt 1.4 vorgestellten Werkzeuge zur automatischen Evaluierung und die dazugehörigen Nutzermodelle hinsichtlich der oben aufgeführten Definition betrachtet, so fällt auf, dass einige Werkzeuge bzw. Methoden keine, im Sinne dieser Arbeit definierten, Nutzermodelle integrieren (HUI Analyzer, GOMS, KLM, ACT-R und CogTool) und dass bei den Werkzeugen, die dies tun, oftmals unterschiedliche Ansätze gewählt wurden.

Nachfolgend werden nun zuerst die Werkzeuge und Methoden kurz vorgestellt, die keine Nutzermodell integrieren (HUI Analyzer, GOMS, KLM, ACT-R und CogTool), bei denen allerdings andere Versuche unternommen wurden, um Vorhersagen bezüglich der Usability für ältere Personen anzupassen. Im Anschluss daran werden die Nutzermodelle von SAMMIE, dem Biswas-Simulator, von DIAS und der MeMo-Werkbank erläutert.

Der HUI Analyzer benötigt, durch die heuristische Evaluation, kein Nutzermodell um zu funktionieren, weshalb das Werkzeug in diesem Abschnitt keine weitere Betrachtung finden wird. Auch KLM und GOMS integrieren keine Nutzermodelle, die der Definition in dieser Arbeit entsprechen, da die Berechnungen der beiden Methoden auf Zeitangaben basieren, die aus empirischen Untersuchungen abgeleitet und schließlich im Modell des Model-Human-Processor (MHP) zusammengefasst wurden [32]. Der MHP beschreibt, wie Informationen während der Interaktion mit einem System wahrgenommen, verarbeitet und schlussendlich in Eingaben umgesetzt werden und definiert hierdurch eine Art *Meta-Nutzermodell*. KLM wie auch GOMS können durch die Anpassung von einzelnen Parametern an Einschränkungen von älteren Personen angepasst werden (siehe z. B. [116]), allerdings

können durch die Art und Weise der Anwendung der Methoden keine demografische Verteilung reproduziert werden, was nach Persad et al. [178] notwendig ist, um Produkte schon in frühen Designphasen effizient validieren zu können.

ACT-R ist eine kognitive Architektur deren Parameter durch Programmierung angepasst werden können. Durch Parameteränderungen ist der Modellierer in der Lage, die Einschränkungen und Eigenschaften von realen Nutzern abzubilden. Ein Problem, welches mit der ACT-R Laufzeitumgebung nur unter großen Schwierigkeiten gelöst werden kann, ist die Wiedergabe von Verteilungen wie sie in der Gesellschaft oder auch nur in Teilen der Gesellschaft vorgefunden werden können. ACT-R simuliert in einem Simulationsdurchgang nur eine spezifische Ausprägung und benötigt entsprechende Anpassungen der Parameter und Wiederholungen der Simulation, um Verteilungen in der Gesellschaft abbilden zu können.

Da das CogTool auf ACT-R aufsetzt sind die genannten Einschränkungen von ACT-R auf dieses Werkzeug zu übertragen. Allerdings wurden im Jahr 2012 von Trewin et al. [234], in einem ersten Versuch, die Parameter des CogTool's angepasst, um die Ausführungszeiten für ältere Versuchspersonen vorherzusagen. In der Auswertung zeigte sich, dass durch die Simulation mit dem CogTool-Explorer signifikant längere Ausführungszeiten für die älteren Personen vorhergesagt wurden, als im realen Test beobachtet werden konnten. Als Begründung für die unterschiedlichen Ausführungszeiten in Vorhersage und empirische Daten wurden fehlende Parameter für das verwendete Eingabegerät genannt. So wurden in diesem Versuch, der einen Computer mit Maus als Eingabegerät nutzte, Parameter verwendet, die aus einem empirischen Versuch mit Mobiltelefonen [117] stammen. Trewin et al. kommen daher zu dem Schluss, dass Vorhersagen über Ausführungszeiten nur möglich sind, wenn die Parameter für das entsprechende Eingabegerät vorliegen. Über eine Weiterentwicklung zur Anpassung des Werkzeugs an die Eigenschaften und Einschränkungen von älteren Personen ist nichts bekannt.

Als nächstes werden die Werkzeuge betrachtet, die Nutzermodelle integrieren. Für jedes Werkzeug wird dargestellt, in welcher Form Wissen über die Eigenschaften und Einschränkungen von Nutzern gespeichert wird und wie dieses Wissen in der Simulation verwendet wird.

Wie in Abschnitt 1.4.1 dargestellt, ist SAMMIE eines der ältesten Werkzeuge um die Interaktion zwischen Mensch und Maschine zu simulieren. SAMMIE nutzt ein 3D-Modell des Menschen um die Sichtbarkeit von Anzeigen und die Erreichbarkeit von Bedienelementen zu simulieren; es nutzt hierfür anthropometrische Daten. Anthropometrische Daten beschreiben die Maße des menschlichen Körpers und werden vor allem dazu genutzt, die Ergonomie von Produkten zu

testen. Um die Körpermaße realistisch wiederzugeben sind in SAMMIE die Daten von über 100 Personen hinterlegt. Wird SAMMIE mit der HADRIAN-Datenbank kombiniert, so stehen für die Simulation weitere 100 Nutzermodelle zur Verfügung, welche Nutzer nachbilden, die zu den Gruppen der älteren oder körperlich eingeschränkten Personen zählen [180]. Die Nutzermodelle in HADRIAN (Human Anthropometric Data Requirements Investigation and ANalysis) bilden reale Nutzer aus empirischen Versuchen ab. Jedes Nutzermodell enthält Angaben über die Körpermaße und die Mobilität einer entsprechenden Versuchsperson. Für einige Nutzermodelle existieren Fotos und Videos, die die zugrundeliegende Versuchsperson zeigen, wie sie Aufgaben löst. Die in SAMMIE bzw. HADRIAN enthaltenen Nutzermodelle ermöglichen dem Entwickler einen vielfältigen Einblick in die Auswirkungen von Einschränkungen, sie ermöglichen allerdings keine Abschätzung der Schwere eines Problems, da die Häufigkeiten von Einschränkungen in den Nutzermodellen keiner demografischen Verteilung folgen.

Der Biswas-Simulator (BS), wie auch DIAS, nutzen XML-basierte Ontologien um Nutzereigenschaften und Nutzereinschränkungen zu speichern. Beide Implementierungen geben für verschiedene Einschränkungen jeweils Hinweise, wie hoch der Anteil in der Bevölkerung ist. In Abbildung 3 und in Abbildung 4 sind Screenshots der beiden Simulatoren dargestellt. Es musste auf Video-Screenshots zurückgegriffen werden, da weder von DIAS noch vom BS eine funktionsfähige Version im Internet gefunden werden konnte. Wie in Abbildung 3 für DIAS dargestellt, wird dem Entwickler für einige Einschränkungen, z. B. *Katarakt*, angezeigt, wie viele Personen die Einschränkung mit dem gewählten Level in der Gesellschaft haben. Die in DIAS verwendeten Daten über die Verteilung in der Gesellschaft wurden alle aus der Literatur übernommen, so resultieren die Angaben über die von einem Katarakt betroffenen Personen aus dem Irish National Disability Survey 2006 [176]. Beim BS kann der Entwickler einstellen, welche Grad von Einschränkung das Nutzermodell haben soll (siehe Abbildung 4). In Abhängigkeit von den ausgewählten Gruppen wird dem Nutzer angezeigt, wie hoch der Anteil der Personen in der Bevölkerung ist, die diese Einschränkung haben. Die Quelle der Daten ist nicht bekannt.

Die MeMo-Werkbank verwendet zur Generierung eines Nutzermodells eine Nutzergruppe. Eine Nutzergruppe beschreibt die Fähigkeiten, Einschränkungen, Stimmungen und Eigenschaften von einer Gruppe von Nutzern. In Tabelle 1 sind die Attribute aufgeführt mit denen eine Nutzergruppe beschrieben werden kann. Die Attribute in der Werkbank können binär sein, wie z. B. für *Blind* (*ja, nein*), mehrere Level haben, wie z. B. *Technikaffinität* (*niedrig, mittel, hoch*), oder einen Zahlenbereich umfassen, wie das *Alter* (z. B. 50 - 75). Die Nutzermodelle werden in der MeMo-Werkbank durch das Zufallsstich-

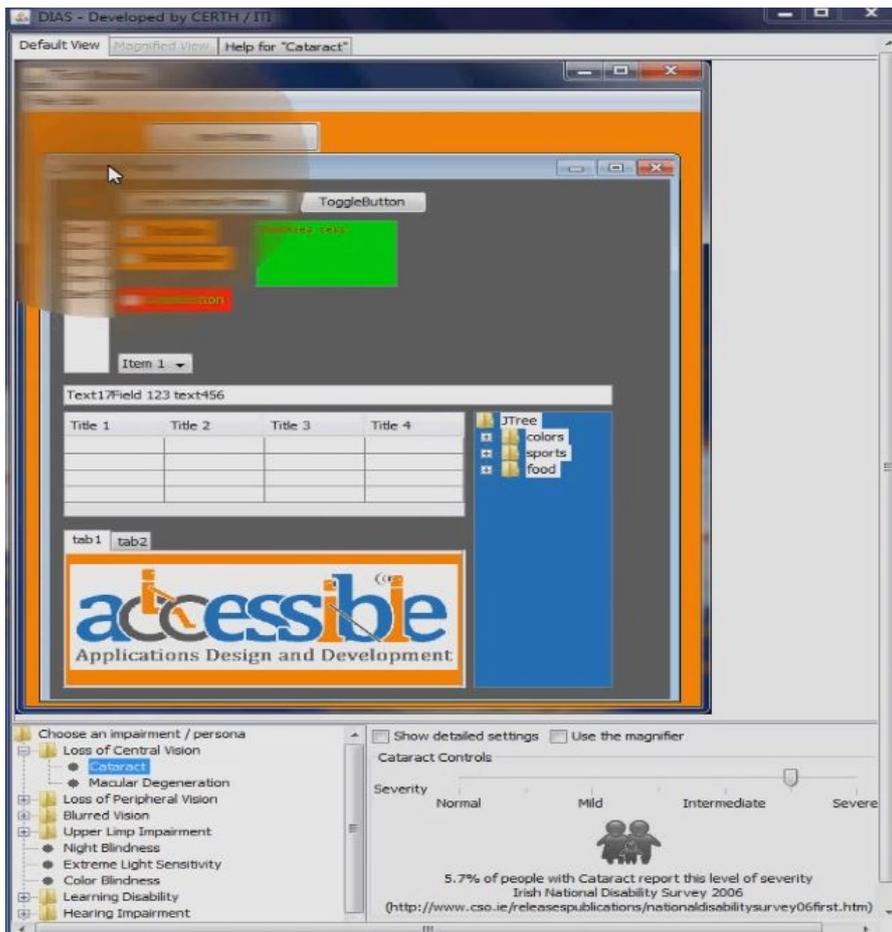


Abbildung 3: Screenshot DIAS

probeverfahren generiert. Für Attribute, die eine binäre Ausprägung oder ein Level haben, ist die Eigenschaft durch die Festlegung des Modellierers definiert – wenn das Nutzergruppenattribute für *Blind* mit *Ja* festgelegt ist, so sind alle generierten Nutzermodelle blind. Für Attribute, die einen Zahlenbereich umfassen, wird durch einen Zufallsalgorithmus eine Ausprägung ermittelt, wobei die Wahrscheinlichkeit für jede mögliche Ausprägung gleich hoch ist. Die Implementierung der Nutzermodelle in der Werkbank nutzt keine demografischen Daten.

Neben den bereits genannten Implementierungen existieren auch andere Methoden, um Informationen über Nutzer zu speichern (z. B. GUMO [93]), allerdings werden diese Implementierungen hier nicht weiter betrachtet, da sie Nutzermodelle im Sinne der zweiten Definition, folglich zur Adaption des Systems an den Nutzer, verwenden.

Werden die vorgestellten Nutzermodelle betrachtet, so fällt auf, dass nur wenige Implementierungen demografische Daten verwendeten (z. B. Biswas-Simulator, DIAS), allerdings haben auch diese Lösungen den Nachteil, dass für den Modellierer nicht in allen Fäl-

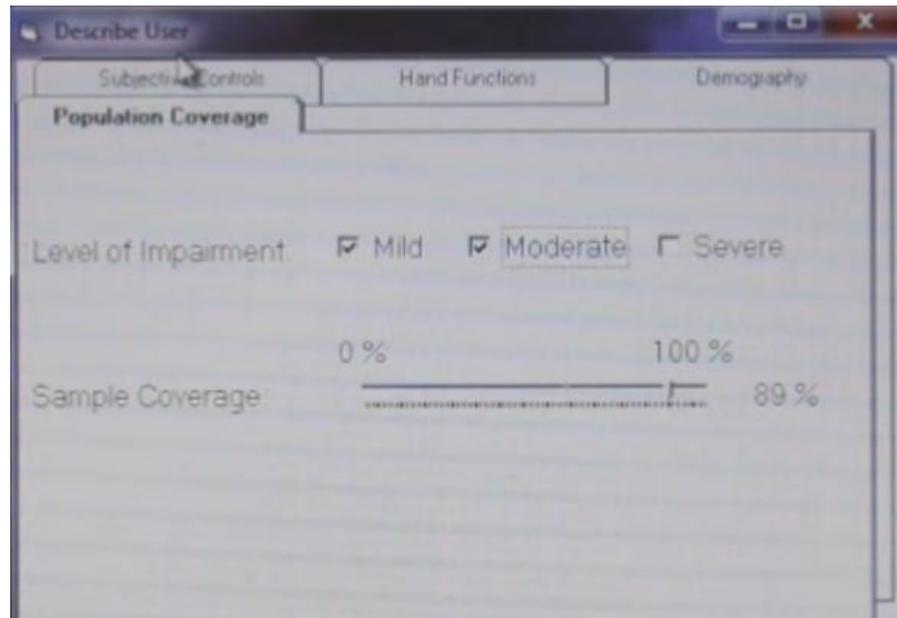


Abbildung 4: Screenshot Biswas-Simulator

len klar ist, wie viele potenzielle Nutzer von einer Einschränkung betroffen sind. Zusätzlich werden Zusammenhänge zwischen verschiedenen Einschränkungen nicht in den Nutzermodellen abgebildet. Eine häufige Begleiterscheinung von Typ 2-Diabetes mellitus, umgangssprachlich auch Altersdiabetes genannt, ist beispielsweise die diabetische Retinopathie – eine Erkrankung der Netzhaut des Auges, die in seltenen Fällen zur Erblindung der Person führen kann, häufig allerdings zu Beeinträchtigungen im Sehen führt. Nach der UK Prospective Diabetes Study [236] haben etwa 36 % der Patienten, bei denen eine Typ 2-Diabetes mellitus erstmalig festgestellt wurde, bereits eine Retinopathie. Kakarlapudi und Kollegen [123] konnten zeigen, dass ein Typ 2-Diabetes mellitus sich ebenfalls auf die Hörfähigkeit auswirkt. Somit kann festgestellt werden, dass visuelle und auditive Einschränkungen nicht unweigerlich unabhängig sind, sondern, in bestimmten Fällen, eine gemeinsame Ursache haben können, die bei der Gestaltung der Benutzeroberfläche beachtet werden müssen. So ist das Hinzufügen eines auditiven Feedbacks nicht zwangsläufig dafür geeignet ein schwaches visuelles Feedback auszugleichen.

1.5 ZIELE UND AUFBAU DER ARBEIT

Werden die bisher vorliegenden Lösungen zur automatisierten Evaluation der Usability betrachtet, so zeigen sich, insbesondere mit Blick auf die Simulation von Interaktionsfehler von Senioren, verschiedene offene Probleme. Einige Werkzeuge bzw. Methoden (z. B. CogTool oder GOMS) sind nur geeignet die Performanz der Nutzer zu simulieren – allerdings ist Performanz, für die Gruppe der älteren Nutzer,

Kategorie	Attribute
Allgemein	Altersspanne
Sprachkenntnisse	Deutsch, Englisch, Französisch, Italienisch, Spanisch
Einschränkungen	Farbenblind, Blind, Stumm, Taub, Tremor, Sehfähigkeit
Psychologische Eigenschaften	Technikaffinität, Vertrauen in Technik, Explorationswille, Computerangst, Gewissenhaftigkeit, Problemlösestrategie
Fähigkeiten	Technische Fähigkeiten, Bildung, Domänenwissen, Aufmerksamkeitsspanne, kognitive Fähigkeiten

Tabelle 1: Attribute der Nutzergruppen

nur ein wichtiger Faktor, neben Sicherheit und Komfort, den es zu evaluieren gilt [67, S. 238]. Ein Werkzeug zur Evaluation von Benutzerschnittstellen für ältere Nutzer soll nicht nur in der Lage sein, die Performanz zu evaluieren, sondern es muss auch Fehler in der Interaktion und deren Auswirkungen auf den Interaktionsprozess simulieren, um eine Abschätzung hinsichtlich der Sicherheit und dem Komfort zu ermöglichen.

Andere Tools, wie DIAS oder der Biswas-Simulator, bieten zwar die Möglichkeit die Einschränkungen der potenziellen Nutzer zu simulieren, allerdings fehlen dort häufig Angaben über die Anzahl der betroffenen Personen und es sind, in den vorliegenden Implementierungen, keine Zusammenhänge zwischen unterschiedlichen Einschränkungen definiert. Nach Persad et al. [178] benötigen Designer umfangreiches Wissen über die späteren Nutzer und deren Einschränkungen, um Produkte schon in frühen Designphasen effizient validieren zu können. Persad und Kollegen empfehlen zur Lösung dieses Problems quantitative Angaben über Verteilungen (Einschränkungen und Leistungsfähigkeit) in der Bevölkerung zu nutzen.

Neben den fehlenden quantitativen Aussagen zu Verteilungen von Einschränkungen in der Gesellschaft, sind auch fehlende oder unpassende mentale Modelle von Eingabegeräten und deren Auswirkungen auf die Interaktion bisher von keinem der vorgestellten Werkzeuge ausreichend abgedeckt. Die Schwierigkeit, die DMM in die automatische Evaluation einzubinden, liegt einerseits in der Natur der mentalen Modelle (siehe hierzu Abschnitt 1.3.3), andererseits auch in der raschen technologischen Entwicklung und den daraus resultierenden notwendigen kontinuierlichen Änderungen, die den potenziellen Nutzer dieser Technologie zum ständigen Lernen neuer Interaktions-

formen nötig [87] und eine kontinuierliche Anpassung der Werkzeuge an den aktuellen Stand der Technik erfordert.

Das Ziel der vorliegenden Arbeit ist die Vorstellung der Ergebnisse zur Lösung der genannten offenen Probleme. Konkret geht es darum zu zeigen, wie vorhandene Daten über Einschränkungen in der Bevölkerung für die automatische Evaluation genutzt werden können. Hierfür wird in Kapitel 3 eine technische Lösung – basierend auf einem Bayes'schen Netz – vorgestellt, die in der Lage ist, aufgrund von beobachteten Daten Vorhersagen zu nicht beobachteten Informationen zu treffen. Das zweite in dieser Arbeit adressierte Problem betrifft die Simulation von unpassenden bzw. fehlenden mentalen Modellen von Eingabegeräten und deren Auswirkungen auf die Interaktion mit einem Smartphone (Kapitel 4). Die dritte Herausforderung ist die Integration der beiden Lösungen in die MeMo-Werkbank und die Validierung der so erhaltenen Gesamtlösung (Kapitel 5).

Diese Arbeit gliedert sich wie folgt. In Kapitel 2 wird die verwendete technische Basis (MeMo-Werkbank) vorgestellt, ferner wird erläutert, wie die Werkbank aufgebaut ist, welche Schritte notwendig sind, um ein Systemmodell zu konstruieren, und wie die bisherigen Ergebnisse der Simulation aussehen. Das Kapitel 3 stellt einen Ansatz dar, wie Daten über die Interaktion beeinflussende Einschränkungen in der Bevölkerung gespeichert und für Vorhersagen genutzt werden können. Der Ansatz, wie unpassende DMM und deren Auswirkungen auf die Interaktion simuliert werden können, ist in Kapitel 4 dargestellt. Daran anschließend werden in Kapitel 5 ein empirischer Versuch und dessen Ergebnisse beschrieben, die dazu dienen, die technischen Lösungen zu validieren. Den Abschluss der Arbeit bildet die Zusammenfassung und der Ausblick in Kapitel 6.

Die MeMo-Werkbank ist ein Werkzeug, um Designideen bzw. Prototypen in frühen Phasen der Entwicklung automatisch hinsichtlich vorhandener Usability-Probleme zu evaluieren [60]. Die Werkbank wurde in der Anfangsphase genutzt, um Modelle von Sprachdialogsystemen (SDS) zu evaluieren [164]. In 2007 konnten Jameson et al. [113] zeigen, dass mit der Werkbank auch grafische Benutzerschnittstellen evaluiert werden können, die mit der Maus bedient werden. Über die vergangenen Jahre wurde die Software stetig weiterentwickelt. So wurde im Jahr 2008 von Feuerstack et al. [66] die *Multi-Access Service Platform* (MASP) mit der MeMo Werkbank kombiniert. Die MASP ist ein System, welches eine multimodale Benutzerschnittstelle auf Basis von abstrakten Modellbeschreibungen generieren kann, die an ein entsprechendes Ausgabegerät angepasst ist [19]. Durch Engelbrecht [59] wurde die Werkbank dahingehend erweitert, dass auf Grundlage der durch die Simulation erzeugten Logdateien, die Qualität von SDS vorhergesagt werden kann.

Die MeMo-Werkbank unterscheidet sich von den in Abschnitt 1.4.1 dargestellten Ansätzen dahingehend, dass die Auswahl des nächsten Interaktionsschrittes durch das Nutzermodell erfolgt und nicht vom Entwickler vorgegeben wird. Die Auswahl des nächsten Interaktionsschrittes erfolgt in der MeMo-Werkbank anhand von Wahrscheinlichkeiten, die durch Nutzermodell- und Bedienelement-Eigenschaften und Regeln beeinflusst werden.

In diesem Kapitel wird erläutert, aus welchen Teilen die MeMo-Werkbank besteht; welche Modelle konstruiert werden müssen, um ein System automatisch evaluieren zu können; wie der Ablauf einer Simulation ist; und was als Ergebnis einer Simulation erwartet werden kann.

2.1 TEILE DER MEMO-WERKBANK

Das Ziel dieses Abschnittes ist es, einen Überblick über die einzelnen Teile der MeMo-Werkbank zu geben und deren Zusammenspiel darzustellen. Hierfür wird im ersten Unterabschnitt dargestellt, wie mithilfe von Wahrscheinlichkeiten Interaktionspfade generiert werden. Im zweiten Abschnitt wird beschrieben, was eine Rule-Engine ist und wie diese verwendet wird, um Wahrscheinlichkeiten zur Auswahl des nächsten Interaktionsschrittes durch Regeln zu verändern. Der dritte Unterabschnitt dient zur Beschreibung der Teile eines Systemmodells und erklärt wie, ein System mit der MeMo-Werkbank modelliert

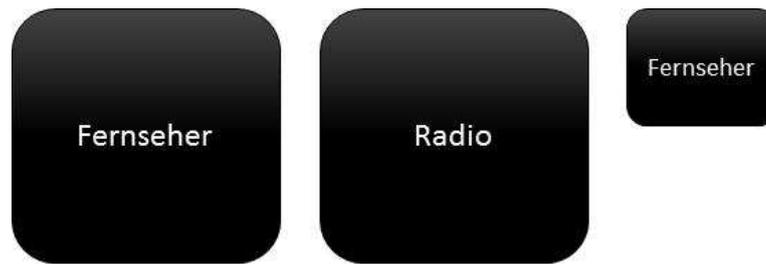


Abbildung 5: Drei unterschiedliche Buttons

werden kann. Nach der Darstellung des Vorgehens zur Systemmodellierung wird das Aufgabenmodell vorgestellt. Im Abschnitt über das Aufgabenmodell wird erläutert, wie Aufgaben in der MeMo-Werkbank modelliert werden können. Im letzten Unterabschnitt wird erläutert, was ein MeMo-Nutzermodell ist, welche Aufgabe es hat, aus welchen Modulen es besteht, und welche verschiedenen Varianten existieren.

2.1.1 Wahrscheinlichkeiten

In der MeMo-Werkbank werden Wahrscheinlichkeiten in der Wahrnehmung und Verarbeitung des Nutzermodells verwendet, um Entscheidungen bezüglich des simulierten Interaktionsverlaufes treffen zu können. Eine Wahrscheinlichkeit (p) repräsentiert hierbei einen Wert, der den Grad der Sicherheit bzw. Gewissheit bezüglich eines Eintretens eines Ereignisses ausdrückt. Eine Wahrscheinlichkeit kann jeden Wert im Intervall zwischen 0 und 1 annehmen, wobei 0 wie auch 1 zulässige Werte sind. Hat ein Ereignis eine Wahrscheinlichkeit von 0, so kann dieses Ereignis unmöglich eintreten. Eine Wahrscheinlichkeit von 1 drückt das sichere Eintreten eines Ereignisses aus. Alle Werte zwischen 0 und 1 definieren die Einschätzung wie sicher ein Ereignis eintreten wird, wobei mit höherer Wahrscheinlichkeit die Einschätzung der Sicherheit des Eintretens steigt.

Die in der MeMo-Werkbank verwendeten Wahrscheinlichkeiten können durch Regeln und Algorithmen verändert werden. Zur Illustration, wie die Wahrscheinlichkeiten die Auswahl eines Interaktionsobjektes beeinflussen, soll das nachfolgende Beispiel dienen. In [Abbildung 5](#) sind drei unterschiedliche Buttons skizziert, die in einem fiktiven System genutzt werden können, um eine Aktion auszulösen. Die beiden Buttons, die mit dem Wort *Fernseher* beschriftet sind, können dazu genutzt werden einen Fernseher einzuschalten. Der Button mit dem Label *Radio* schaltet das Radio an. Die Aufgabenstellung des Nutzers ist es, den Fernseher anzuschalten. Für das Beispiel wird nur der Verarbeitungs- und nicht der Wahrnehmungsprozess betrachtet,

die Verwendung von Wahrscheinlichkeiten ist gleichwohl in beiden Prozessen äquivalent.

Ist der Verarbeitungsprozess angestoßen, wird zuerst die Anzahl der verfügbaren Interaktionsoptionen vom Verarbeitungsmodul der MeMo-Werkbank bestimmt. Im fiktiven Beispiel existieren 3 Interaktionsoptionen. Da nach diesem ersten Evaluationsschritt noch nicht bestimmt werden kann, welches Objekt verwendet werden soll, haben alle Interaktionsoptionen die gleiche Wahrscheinlichkeit ($p = 0,333$).

In einem nächsten Schritt wird jeder Button bezüglich der Relevanz für die Aufgabenstellung bewertet. Werden bei diesem Vergleich eine oder mehrere Übereinstimmungen gefunden, so werden die Wahrscheinlichkeiten der entsprechenden Interaktionsoptionen erhöht. Der Wert, um den der Wahrscheinlichkeitswert einer Interaktionsoption erhöht wird, hängt von der Anzahl der gefundenen relevanten Optionen und von dem in der Werkbank definierten Wert für relevante Interaktionsoptionen ab. Der Wahrscheinlichkeitswert für relevante Interaktionsoptionen kann vom Entwickler festgelegt werden. Wenn der Wert nicht vom Entwickler festgelegt wird, so wird als Wahrscheinlichkeitswert der Wert $p = 0,95$ verwendet. Im genannten Beispiel passen 2 Buttons (der große und der kleine Fernsehbuttons) zur Aufgabenstellung. Da nicht beide Fernsehbuttons eine Wahrscheinlichkeit von 0,95 haben können muss die Wahrscheinlichkeit zwischen beiden Buttons aufgeteilt werden. Als Resultat des Vergleiches liegt nun folgende Wahrscheinlichkeitsverteilung vor: Die beiden Fernsehbuttons haben jeweils eine Wahrscheinlichkeit von 0,475. Der Radiobutton, welche nicht zur Aufgabestellung passt, hat nach diesem Schritt eine Wahrscheinlichkeit von 0,05.

Wird angenommen, dass eine Regel existiert, welche aussagt, dass große Button doppelt so häufig verwendet werden wie kleine Button, dann ändern sich die Verteilung der Wahrscheinlichkeiten wie folgt: Nach Anwendung der Regel hat der große Fernsehbutton einen Wahrscheinlichkeitswert von 0,633. Der kleine Fernsehbutton hat einen Wahrscheinlichkeitswert von 0,316 und der Radiobutton behält den Wert 0,05. Würde ein kleiner Radiobutton existieren, so hätte sich die Restwahrscheinlichkeit von 0,05 auf 0,033 für den großen Radiobutton und 0,016 für den kleinen Radiobutton aufgeteilt.

Nachdem die Wahrscheinlichkeiten der Button berechnet wurden, wird durch einen Algorithmus eine Interaktionsentscheidung herbeigeführt. Die Berechnung der Interaktionsentscheidung erfolgt wie folgt: Im ersten Schritt wird eine zufällige Zahl zwischen 0 und 1 erzeugt, wobei eine Gleichverteilung angenommen wird. Für das genannte Beispiel wird davon ausgegangen, dass die erzeugte Zufallszahl den Wert 0,7 hat. Um zu entscheiden welche Option gewählt wird, werden die Wahrscheinlichkeitswerte der Interaktionsoptionen (Button) addiert bis die Summe der Werte größer oder gleich der Zufallszahl ist. Ist die Summe größer oder gleich der Zufallszahl, so wird die



Abbildung 6: Schematische Darstellung einer Regel in der MeMo-Werkbank

aktuelle Interaktionsoption als Interaktionsentscheidung gespeichert. Die Reihenfolge der Addition ist, aufgrund der Gleichverteilung bei der Erzeugung der Zufallszahl, nicht entscheidend, auch wenn dies für den konkreten Fall anders ist. Für das verwendete Beispiel sei angenommen, dass die Buttons in der Reihenfolge großer Fernsehbutton, Radiobutton und kleiner Fernsehbutton addiert werden. Da die Summe der Addition der Wahrscheinlichkeitswerte des großen Fernsehbuttons und des Radiobuttons kleiner als 0,7 sind, wird für diesen Fall die letzte Interaktionsoption (kleiner Fernsehbutton) als Interaktionsentscheidung gespeichert.

2.1.2 Regeln und Regelverarbeitung

Eine in der MeMo-Werkbank verwendete Regel beschreibt, als Wenn-Dann-Aussage, das Wissen über die Interaktion beeinflussende Eigenschaften von Bedienelementen und deren Konsequenzen auf den Interaktionsprozess. Die Regeln dienen in der Werkbank dazu, die Wahrscheinlichkeiten der Interaktionsoptionen zu beeinflussen. Zur Verarbeitung der Regeln während des Interaktionsprozesses wird eine Rule-Engine verwendet.

Die Regeln, die in der MeMo-Werkbank verwendet werden, beschreiben Erkenntnisse aus Usability-Versuchen, der HCI-Literatur und Usability-Heuristiken als Wenn-Dann-Aussagen und sind in einem XML-Format gespeichert. In Abbildung 6 ist der Aufbau einer MeMo-Regel schematisch dargestellt. Eine MeMo-Regel besteht aus einem Deklarationsteil, einem Bedingungsteil und einem Teil, der die Konsequenz beschreibt.

Der Deklarationsteil definiert den Namen der Regel, zusätzlich kann optional eine Beschreibung eingefügt werden, die die Herkunft

der Regel erklärt. Durch die Definition des Bedienelementes (z. B. *Button*) in der Deklaration wird festgelegt, für welche Elemente die Regel angewendet werden kann. Werden zusätzlich Attribute mit den entsprechenden Werten angegeben (z. B. *Schriftgröße = klein*), so wird die Anwendbarkeit auf Bedienelemente eingeschränkt, die die entsprechenden Eigenschaften besitzen (*Button mit kleiner Schriftgröße*). Der zweite Teil einer Regel, der Bedingungsteil, ist optional. Im Bedingungsteil einer Regel können weitere Einschränkungen für die Anwendbarkeit definiert werden. Der Bedingungsteil kann Nutzerattribute (*Nutzer älter als 70 Jahre*) als Einschränkungen enthalten, wodurch die Anwendbarkeit einer Regel weiter differenziert werden kann (z. B. *Button mit kleiner Schriftgröße und Nutzer älter als 70 Jahre*). Damit die Konsequenz einer Regel ausgeführt wird, müssen sowohl die Voraussetzungen der Deklaration als auch die der zusätzlichen Bedingungen erfüllt werden. Im Konsequenz-Teil der Regel ist beschrieben, wie sich eine Wahrscheinlichkeit verändert. Da pro Bedienelement mehrere Regeln angewendet werden können, sind die Veränderungen der Wahrscheinlichkeit im Konsequenz-Teil über Symbole kodiert. Soll die Konsequenz einer Regel die Wahrscheinlichkeit verringern, so kann dies durch {-} ausgedrückt werden. Soll die Wahrscheinlichkeit durch eine Regel erhöht werden wird {+} verwendet. Die maximale Verringerung kann durch 3 Minus und die maximale Erhöhung durch 3 Plus ausgedrückt werden. In der MeMo-Werkbank erfolgt die Verarbeitung der Regeln mit Hilfe einer Rule-Engine.

Die in der MeMo-Werkbank verwendete Rule-Engine wurde von Friedmann-Hill [69] entwickelt und besteht aus drei Teilen; einer Inferenz-Maschine, einer Regelbasis, und einem Arbeitsspeicher. Nach Darstellung von Friedmann-Hill ist die Inferenz-Maschine das wichtigste Element der Rule-Engine, da dieser Teil die Anwendung der Regeln steuert. Die Inferenz-Maschine besteht wiederum aus 3 Bestandteilen; einer Mustererkennung, einer Agenda und einem Programm zur Ausführung des Dann-Teils einer Regel. In der Inferenz-Maschine werden, durch einen Mustervergleich, die Regeln aktiviert, deren Wenn-Teil zu den im Arbeitsspeicher der Rule-Engine gespeicherten Fakten passen. In der MeMo-Werkbank beschreiben Fakten die Eigenschaften von Bedienelementen in der aktuellen modellierten Benutzerschnittstelle. Werden Regeln aktiviert, so werden sie von der Inferenz-Maschine auf die Agenda geschrieben. Die Agenda definiert dabei die Reihenfolge, in welcher die aktivierten Regeln ausgeführt werden. Sind alle Regeln durch den Mustervergleich überprüft worden, so werden, durch den Ausführungsteil der Inferenz-Maschine, die Dann-Teile der aktiven Regeln angewendet. Da eine Regel die Wahrscheinlichkeit einer Interaktionsoption verändert, werden durch die Anwendung der Dann-Teile die Wahrscheinlichkeiten entweder erhöht oder verringert.

2.1.3 Systemmodell

Das Systemmodell ist ein modellierter Prototyp einer Designidee. Es stellt somit keine abgeschlossene Lösung dar, sondern einen Zwischenstand, welche iterativ verbessert werden soll. Ein Systemmodell wird anstelle eines realen Prototyps konstruiert, da die Modellierung des Systemmodells in der MeMo-Werkbank im Allgemeinen schneller ist als die Realisierung eines funktionierenden Prototyps. Ein Systemmodell muss vom Entwickler erstellt werden. Für die Konstruktion des Systemmodells stehen dem Entwickler zwei unterstützende Werkzeuge, der Dialogdesigner und der Systemdesigner, in der Werkbank zur Verfügung. Ist das Systemmodell erstellt, wird es von der MeMo-Werkbank durch Simulation von Interaktionspfaden evaluiert.

2.1.3.1 Dialogdesigner

Der *Dialogdesigner* unterstützt den Entwickler bei der Modellierung der Benutzerschnittstelle. Die Benutzerschnittstelle ist der Teil vom System, der vom Nutzer bzw. vom Nutzermodell wahrgenommen werden kann und der die Interaktion mit dem System ermöglicht. Eine Benutzerschnittstelle in der MeMo-Werkbank kann visuell sein, die Schnittstelle zu einem Sprachdialogsystem (SDS) nachbilden oder beide Modalitäten kombinieren.

Die Abbildung 7 zeigt die Oberfläche des Dialogdesigners. Der Dialogdesigner verwaltet den Dialogpool des Systemmodells. Der Dialogpool enthält die vom Entwickler konstruierten Dialoge, welche vom Nutzermodell zur Interaktion mit dem System genutzt werden können. Ein Dialog ist ein zusammenhängender Teil einer Benutzerschnittstelle. In Abbildung 7 ist der Dialog mit dem Namen *Start* ausgewählt und auf der rechten Seite abgebildet. Der Dialog *Start* enthält ein Bild (*start.png*), welches die Benutzeroberfläche darstellt. Weiterhin enthält der ausgewählte Dialog verschiedene Gruppen von Bedienelementen (*Top*, *Main*, *Zusatz* und *Unten*). Im gewählten Beispiel bestehen die Gruppen aus Sammlungen von Buttons, eine Gruppe kann allerdings auch andere Bedienelemente enthalten. In der Gruppe *Main* ist der Button mit der Bezeichnung *Fernseher* selektiert. Durch die Auswahl werden die Eigenschaften des Buttons auf der rechten unteren Seite angezeigt. Wie in der Abbildung dargestellt, sind für den Button *Fernseher* nur wenige Attribute mit Werten belegt. So wurde für den Button modelliert, dass er ein Label besitzt, welches den Text *Fernseher* enthält. Weiterhin wurde für den Button definiert, dass die Größe *groß* ist, dass er *nicht hervorgehoben* ist und keine Navigationseigenschaften im Sinne von vorwärts oder rückwärts hat.

In Tabelle 2 sind die Eigenschaften dargestellt, die für grafische Eingabelemente definiert werden können. Hat ein Attribut eine binäre Ausprägung, so können diesem Attribute nur die Werte *ja* oder *nein* zugeordnet werden. Hat ein Attribut eine kategoriale Ausprägung,

Kategorie	Attribut	Ausprägung	Beispiel
Allgemeine Attribute	Name	Text	Fernseher
	Informationen	-	Left Click = selektieren
Visuelle Attribute	Position	Koordinaten	X:25 Y:73
	Höhe	Zahl	122
	Breite	Zahl	122
	Sichtbarkeit	Binär	ja
Abstrakte Attribute	Navigation vorwärts	Binär	nein
	Navigation rückwärts	Binär	nein
	Hervorgehoben	Binär	nein
Spezifische Attribute	Label	Text	Fernseher
	Schatten	Niveau	nicht spezifiziert
	Kontrast Label	Niveau	nicht spezifiziert
	Kontrast	Niveau	nicht spezifiziert
	Schriftgröße	Niveau	nicht spezifiziert
	Sprache	Niveau	nicht spezifiziert
	Fremdwörter	Niveau	nicht spezifiziert
	Hat Label	Niveau	nicht spezifiziert
	Hat Symbol	Niveau	nicht spezifiziert
	Hat Icon	Niveau	nicht spezifiziert
	Ist dreidimensional	Niveau	nicht spezifiziert
	Größe	Niveau	groß
Layoutgruppe	Niveau	nicht spezifiziert	

Tabelle 2: Eigenschaften von visuellen Eingabeelementen am Beispiel des Buttons *Fernseher*

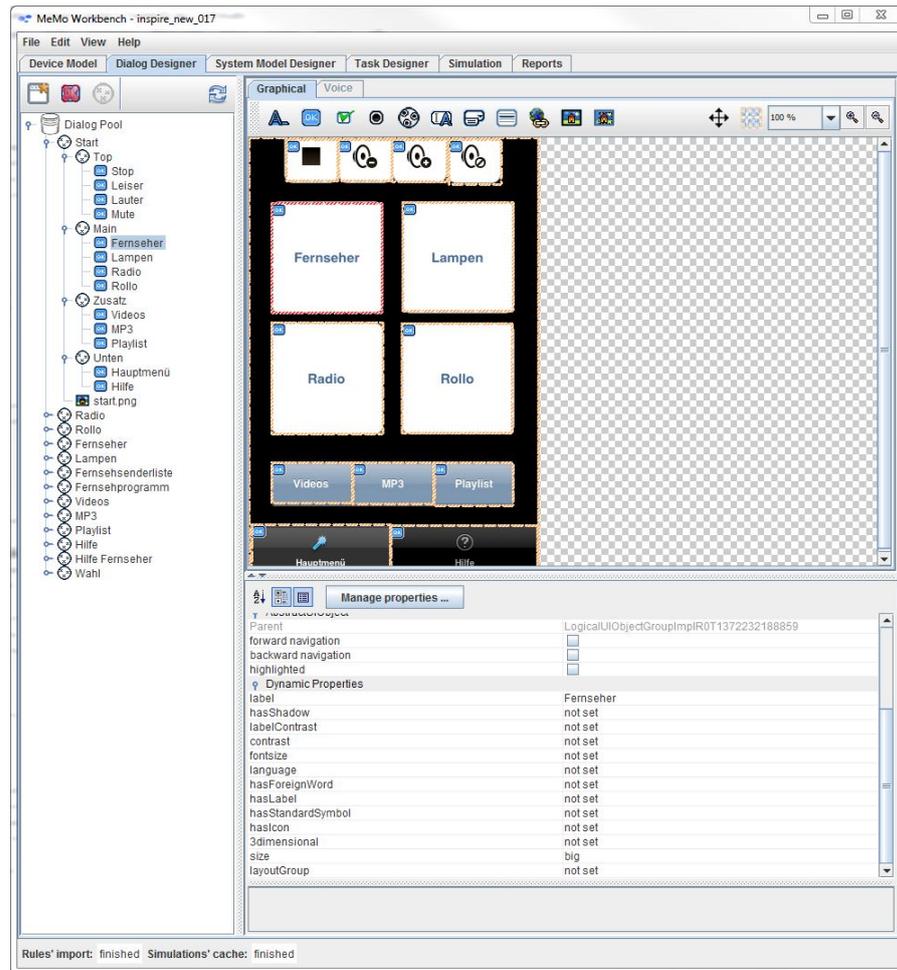


Abbildung 7: MeMo-Werkbank Dialogdesigner

so kann für dieses Element eine entsprechende Ausprägung oder der Wert *nicht gesetzt* (engl. *not set*) ausgewählt werden. Ist der Wert als *nicht gesetzt* definiert, so wird die Eigenschaft während der Simulation keinen Einfluss haben. In der Spalte *Ausprägung* wurde für die Zeile *Informationen* kein Wert (-) definiert, weil die Modellierung von Informationen über den Interaktion-Information-Editor erfolgt, der in Abbildung 8 dargestellt ist. Bevor die Funktion des Interaktion-Information-Editors beschrieben wird, werden die beiden Begriffe *Information* und *Interaktion* näher erläutert.

Eine *Information* wird über ein Attribut-Wert-Paar abgebildet. Das Attribut einer Information ist ein Name. Der Name dient dazu die Information eindeutig zu identifizieren. Der Wert einer Information ist von einem bestimmten Typ (z. B. Text oder Zahl) und kann nur bestimmte, vom Entwickler zu definierende, Ausprägungen annehmen. Die Verwendung von Informationen im Interaktion-Information-Editor dient zur Modellierung von implizitem Wissen eines Nutzermodells. Der zweite Begriff *Interaktion* beschreibt eine Möglichkeit, wie Informationen zwischen Nutzermodell und Systemmodell ausge-

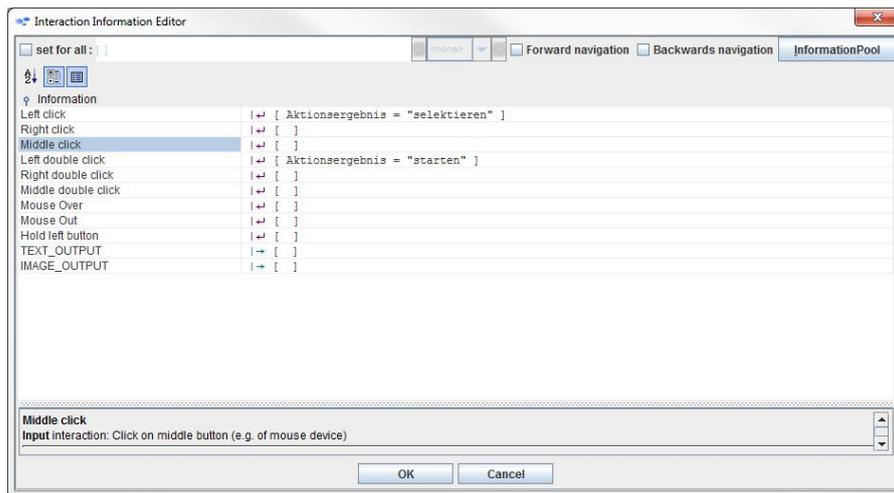


Abbildung 8: Editor für Eingabe- und Ausgabe-Informationen von Eingabe-elementen

tauscht werden können. Es existieren Input-Interaktionen und Output-Interaktionen, wobei Input-Interaktionen den Informationsfluss von Nutzermodell zum Systemmodell und Output-Interaktionen den Informationsfluss vom Systemmodell zum Nutzermodell beschreiben. Zur einfacheren Erläuterung der beiden Konzepte (Information und Interaktion) soll das nachfolgende Beispiel genutzt werden.

Für das Beispiel sei angenommen, dass eine Information mit dem Namen *Aktionsergebnis* vom Typ *Text* angelegt wurde, die als Ausprägungen *selektieren* und *starten* hat. Für den in Abbildung 8 dargestellt Interaktion-Information-Editor bedeutet dies, dass von einem Entwickler modelliert werden kann, dass das Nutzermodell weiß, dass die Aktion *Left click* als Ergebnis etwas *selektiert*. Um dieses Wissen zu modellieren muss der Entwickler als Information für die Input-Interaktion *Left click* angegeben, dass das *Aktionsergebnis* nach Anwendung der Interaktion *selektieren* ist. Nach dem gleichen Prinzip kann definiert werden, dass die Aktion *Double left click* etwas *startet*. Hat das Nutzermodell nun die Aufgabe ein Interaktionselement zu selektieren, so kann es aufgrund der modellierten Information entscheiden, dass die Interaktion (*Left click*) die richtige Interaktion ist, um ein Element zu selektieren.

Um einen Dialog zu modellieren, muss der Entwickler im ersten Schritt einen Dialog, durch Betätigung eines Buttons, erzeugen und den Typ des Dialoges definieren. Durch die Definition des Typs wird festgelegt, ob es sich um einen grafischen oder per Sprache gesteuerten Dialog handelt. Ist der Typ des Dialogs bestimmt, kann der Entwickler verschiedene Elemente zur Benutzerschnittstelle hinzufügen. Wenn Skizzen oder andere Darstellungen existieren, so können diese als Hintergrundbild für grafische Dialoge eingefügt werden. Die Hintergrundbilder sind sinnvoll, da sie die Modellierung und Positionierung der Bedienelemente vereinfacht. Die Positionierung und Definition

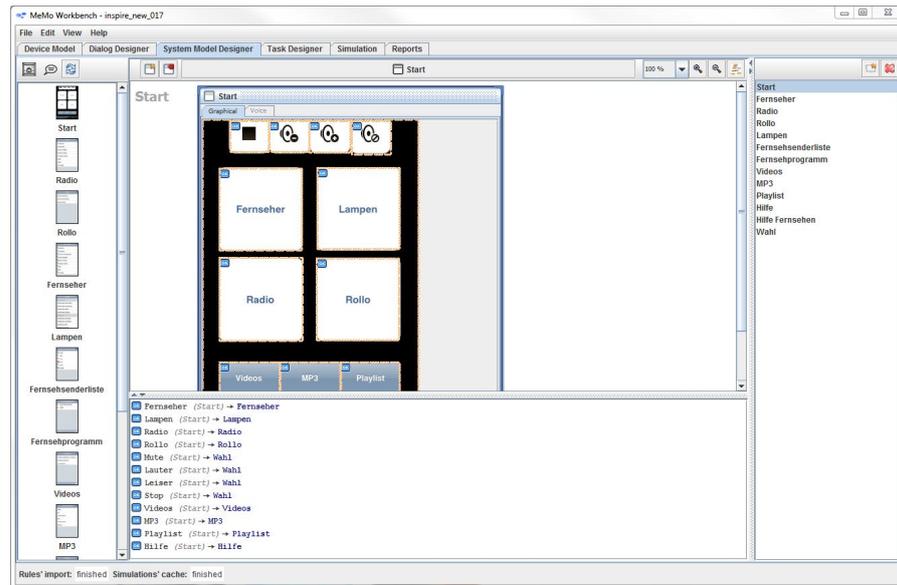


Abbildung 9: Der Systemdesigner der MeMo-Werkbank

der Größe eines Eingabelements erfolgt, indem der Entwickler mit der Maus ein Rechteck um den entsprechenden Bereich zieht. Ist der Bereich eines Elementes definiert muss festgelegt werden, welche Eigenschaften das Element hat und welche Informationen für die möglichen Interaktionen vorliegen. Sind alle Dialoge modelliert, werden sie unter Zuhilfenahme des Systemdesigners in das Systemmodell integriert.

2.1.3.2 Systemdesigner

Der in Abbildung 9 dargestellte *Systemdesigner* dient dazu, verschiedene Dialoge miteinander zu verbinden und so das Verhalten des Systems zu modellieren. Auf der linken Seite der Abbildung sind die verfügbaren Dialoge dargestellt; die rechte Seite der Abbildung zeigt die schon angelegten Systemzustände an, und in der Mitte wird der aktuell ausgewählte Systemzustand mit seinen enthaltenen Dialogen und die zugehörigen Transitionen angezeigt.

Die Systemlogik wird über verschiedene Systemzustände und durch Übergänge zwischen den Systemzuständen, den sogenannten Transitionen, abgebildet. Um einen neuen Systemzustand zu modellieren, muss ein neuer leerer Zustand erzeugt werden. Ist der neue Systemzustand erzeugt, kann er mit einem entsprechenden Namen versehen werden und es können verschiedene Dialoge dem Systemzustand zugeordnet werden. Die Möglichkeit einem Systemzustand mehrere Dialoge unterzuordnen erlaubt es, die Benutzerschnittstelle modular aufzubauen und Dialoge wiederholt zu verwenden.

Ist mindestens ein Dialog dem Systemzustand zugeordnet und hat dieser Dialog mindestens ein Bedienelement, so kann dieses Be-

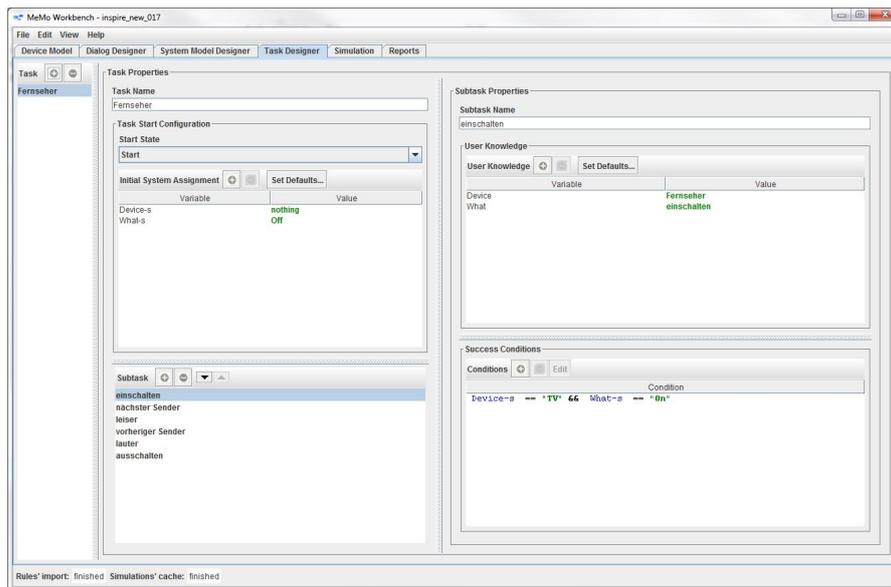


Abbildung 10: Der Aufgabendesigner der MeMo-Werkbank

dienelement mit anderen Systemzuständen über Transitionen, durch einfaches Drag und Drop, verbunden werden. Eine Transition benötigt eine Interaktion; sie verbindet immer genau zwei Systemzuständen und hat *Bedingungen* und *Konsequenzen*. Eine Interaktion kann mehrere Transitionen haben, jedoch müssen die Bedingungen der Transitionen disjunkt sein. Die Bedingungen einer Transition definieren, wann sie ausgeführt werden kann. Durch unterschiedliche Bedingungen kann ein und dieselbe Interaktion zu unterschiedlichen Systemzuständen führen. Jede Transition kann mehrere Konsequenzen haben. Eine Konsequenz verändert eine Information bzw. die Ausprägung einer Information im System- oder Nutzermodell. Durch die Änderung der Informationen kann von der MeMo-Werkbank entschieden werden, wann eine Aufgabe vom Nutzermodell erfüllt wurde.

2.1.4 Aufgabenmodell

Das Aufgabenmodell beschreibt die Aufgabe, die vom Nutzermodell mit dem Systemmodell gelöst werden soll. In Abbildung 10 ist der Aufgabendesigner dargestellt, mit dessen Hilfe der Entwickler das Aufgabenmodell definieren kann. Da ein Systemmodell genutzt werden kann um mehrere Aufgaben zu lösen, bietet der Aufgabendesigner an, dass mehrere Aufgaben definiert werden können. In der Abbildung 10 werden die bereits modellierten Aufgaben auf der linken Seite dargestellt, für das gezeigte Beispiel existiert nur eine Aufgabenstellung mit dem Namen *Fernseher*. Rechts neben der Liste ist der Aufgabeneditor dargestellt. Der Aufgabeneditor erlaubt dem Entwickler für eine Aufgabe einen Namen, einen Startzustand, eine initiale Belegungen von

Informationen, Teilaufgaben, Aufgabenwissen und Zielzustände zu definieren.

Um eine Aufgabe zu modellieren, muss der Entwickler der Aufgabenstellungen einen eindeutigen Namen (z. B. *Fernseher bedienen*) geben. Weiterhin muss er den Startzustand und die initiale Belegung der benötigten Attribut-Wert-Paare definieren. Jede Aufgabe benötigt mindestens eine Unteraufgabe, da nur Unteraufgaben dazu genutzt werden können, das nötige Aufgabenwissen und die Erfolgsbedingungen zu definieren.

Ein Aufgabenmodell soll im Idealfall auf realen Aufgabenstellungen basieren, allerdings werden die Aufgaben in einem Aufgabenmodell nicht textuell beschrieben, sondern als Attribut-Wert-Paare modelliert. Lautet die Aufgabe für den realen Nutzer: „*Schalten Sie den Fernseher an.*“, so kann die Aufgabenstellung für das Nutzermodell in Form zweier Attribut-Wert-Paare als *Gerät = Fernseher* und *Aufgabe = einschalten* modelliert werden. Damit die Werkbank entscheiden kann, ob eine Unteraufgabe abgeschlossen ist, müssen vom Modellierer die Bedingungen definiert werden, die festlegen, wann die Aufgabe erfolgreich erfüllt ist. Sind alle Unteraufgaben mit Erfolg beendet worden, so ist auch die Hauptaufgabe abgeschlossen.

2.1.5 Nutzermodell

Ein Nutzermodell der MeMo-Werkbank repräsentiert einen realen Nutzer des späteren Systems. Es entspricht somit dem ersten Typ von Nutzermodellen, welche in Abschnitt 1.4.2 beschrieben wurden. Das MeMo-Nutzermodell wird in der Werkbank während der Simulation verwendet, um mit dem Modell des Systems zu interagieren. Jedes von der MeMo-Werkbank erzeugte Nutzermodell hat Eigenschaften, die sich aus der Beschreibung der ausgewählten Nutzermodellgruppe ableiten. In der dieser Arbeit zugrunde liegenden Werkbank gibt es 9 verschiedene Nutzermodellgruppen. Jede Nutzermodellgruppe hat Attribute mit denen Sie beschrieben wird. In Tabelle 3 sind – am Beispiel der Nutzermodellgruppe mit dem Namen *Mitte 30, Farbenblind* – die Attribute aufgeführt, die zur Definition einer solchen Nutzermodellgruppe verwendet werden können.

In der MeMo-Werkbank haben Attribute zur Beschreibung von Nutzermodellklassen eine binäre, kategoriale, oder intervallskalierte Ausprägung – die Ausnahme von dieser Regel bildet das Attribut *Name*, welches einen Text enthält. Wenn aus der Beschreibung der Nutzermodellgruppe ein einzelnes Nutzermodell abgeleitet wird, so richtet sich die konkrete Ausprägung der Eigenschaften des Nutzermodells nach den Attributen der Nutzermodellgruppe. Für binäre und kategoriale Attribute der Nutzermodellgruppe wird die gewählte Ausprägung für das Nutzermodell übernommen. Für das Beispiel in Tabelle 3 bedeutet dies, dass alle abgeleiteten Nutzermodelle *Farbenblind*

Kategorie	Attribut	Ausprägung	Beispiel
Allgemeine Attribute	Name	Text	Mitte 30, Farbenblind
	Alter	Intervall	30 - 40
Sprachkenntnisse	Deutsch	Binär	ja
	Englisch	Binär	nein
	Französisch	Binär	nein
	Italienisch	Binär	nein
	Spanisch	Binär	nein
Einschränkungen	Farbenblind	Binär	ja
	Blind	Binär	nein
	Taub	Binär	nein
	Stumm	Binär	nein
	Tremor	Binär	nein
Psychologische Attribute	Sehen	Niveau	Gut
	Technikaffinität	Niveau	hoch
	Vertrauen in Technik	Niveau	Standard
	Wille zur Exploration	Niveau	hoch
	Ängstlichkeit	Niveau	niedrig
	Gewissenhaftigkeit	Niveau	Standard
Fähigkeiten	Problemlösestrategie	Niveau	Analytisch
	Technische Fähigkeiten	Intervall	5 - 10
	Bildung	Intervall	5 - 10
	Domänenwissen	Niveau	Standard
	Aufmerksamkeitsspanne	Niveau	Standard
Dynamische Nutzerattribute	Kognitive Fähigkeiten	Niveau	hoch
	Aufmerksamkeit	Intervall	50 - 100
	Irritation	Intervall	0 - 100
	Interesse	Intervall	40 - 100
	Zeitdruck	Intervall	50 - 100

Tabelle 3: Attribute der Nutzermodellklassen am Beispiel der Nutzermodellklasse *Mitte 30, Farbenblind*

sind und eine geringe Ängstlichkeit aufweisen, weil die Nutzermodellgruppe als Farbenblind und wenig ängstlich definiert wurde. Ist das Attribut der Nutzermodellklasse intervallskaliert, so wird dem Nutzermodell ein zufälliger Wert aus dem Intervall zugewiesen. Die zufällige Zuweisung eines Wertes basiert auf der Annahme, dass alle Werte im Intervall die gleiche Wahrscheinlichkeit aufweisen. Für das dargestellte Beispiel bedeutet dies, dass von 100 erzeugten Nutzermodellen im Mittel etwa 9 Nutzermodelle ein Alter von 30 Jahren haben, da bei 11 möglichen Ausprägungen ($Alter := \{x \in \mathbb{N} \mid 30 \leq x \leq 40\}$), die Wahrscheinlichkeit für eine Ausprägung rund 9 % beträgt.

Jedes durch die MeMo-Werkbank erzeugte Nutzermodell hat nicht nur die aus der Nutzermodellgruppe abgeleiteten Eigenschaften, es hat auch die Fähigkeit Informationen wahrzunehmen, zu verarbeiten und – basierend auf dem Ergebnis der Informationsverarbeitung – das Verhalten des Systemmodells durch Eingaben zu beeinflussen.

Die beschriebene Dreiteilung der Informationsverarbeitung basiert auf dem, von Card 1981 [28] vorgestellten, Model Human Processor (MHP). Der MHP beschreibt die Unterteilung der Informationsverarbeitung in den drei Teilen *Wahrnehmung*, *Verarbeitung* und *Ausführung*. In der MeMo-Werkbank werden die drei Teile der Informationsverarbeitung durch austauschbare Module repräsentiert. Der Aufbau und die Funktionsweise der einzelnen Module für die Wahrnehmung, Verarbeitung und Ausführung werden in den nachfolgenden Unterabschnitten näher erläutert.

2.1.5.1 *Wahrnehmung*

Die Module für die Wahrnehmung werden verwendet, um die vom System bereitgestellten Informationen bzw. Ausgaben wahrzunehmen, und sie an die nachfolgende Verarbeitung zu übergeben.

Die Module zur Wahrnehmung der Informationen bzw. Systemausgaben funktionieren nur mit modellierten Angaben. Es werden keine Bild- oder Texterkennungsalgorithmen eingesetzt um Bedienelemente oder Label in grafischen Benutzerschnittstellen zu erkennen. Der Verzicht auf solche Algorithmen resultiert aus dem Ziel, die MeMo-Werkbank bereits in frühen Phasen der Entwicklung einsetzen zu können, in denen die Gestaltung der Oberfläche noch nicht abgeschlossen ist.

Für die Wahrnehmung der Informationen und Ausgaben existieren zwei verschiedene Module, ein *Default Perception Module* (DPERM) und ein *Bottom-Up Perception Module* (BUPM). Das DPERM hat die Aufgabe eine fehlerfreie Wahrnehmung der durch die Benutzerschnittstelle präsentierten Informationen zu gewährleisten. Diese Funktionalität ist insofern sinnvoll, da sie ermöglicht, Fehler durch eine fehlerbehaftete Wahrnehmung auszuschließen. Um Fehler zu vermeiden, die durch die Reihenfolge der Wahrnehmung entstehen können, nimmt dieses Modul alle Informationen und Ausgaben der Benutzerschnitt-

stelle gleichzeitig wahr und sendet die so gesammelten Informationen gebündelt an das nachfolgende Modul zur Verarbeitung der wahrgenommenen Informationen. Das BUPM, das zweite Modul zur Wahrnehmung von Informationen und Systemausgaben, wählt in einem ersten Schritt die Elemente der Benutzerschnittstelle aus, die besonders salient (auffallend) sind. Sind die Elemente anhand ihrer Salienz sortiert, so werden sie, basierend auf der Reihenfolge, vom Wahrnehmungsmodul verarbeitet. Hat das Modul eine definierte Anzahl von Informationen wahrgenommen, so werden diese zur Auswertung an das Verarbeitungsmodul gesendet. Die Anzahl der durch das BUPM in einem Schritt wahrgenommenen Informationen berechnet sich aus dem maximalen Winkel für den Bereich des schärfsten Sehens (Fovea centralis), der Auflösung des Displays und aus der Entfernung des Nutzers vom Display. Da in der aktuellen Implementierung weder die Auflösung des Displays, noch der Winkel der Fovea centralis oder die Entfernung vom Bildschirm angegeben werden können, werden die nachfolgend dargestellten, aus der Literatur abgeleiteten, Werte verwendet [196]: Blickwinkel 2° , Bildschirmauflösung 100 dpi und Entfernung vom Bildschirm 60 cm. Hat das BUPM die wahrgenommenen Informationen an das Modul zur Verarbeitung der Informationen weitergeleitet, so pausiert es, bis vom Verarbeitungsmodul die Anforderung kommt, dass mehr Informationen benötigt werden.

2.1.5.2 *Verarbeitung*

Die verschiedenen Module der Verarbeitung dienen dazu, anhand der wahrgenommenen Informationen und des definierten Aufgabenwissens zu entscheiden, welche Interaktion vom Nutzermodell auf das Systemmodell angewendet werden soll.

Die Art und Weise wie eine Interaktion ausgewählt wird hängt stark vom jeweiligen Verarbeitungsmodul ab. Im Default Processing Module (DPROM) werden, in einem ersten Schritt, die durch das Wahrnehmungsmodul übermittelten Interaktionsoptionen und die zugehörigen Informationen mit dem Wissen des Nutzermodells verglichen. Wird durch diesen Vergleich eine Interaktionsoption gefunden, die gut zum Nutzerwissen passt, so wird diese Interaktionsoption in einer Liste gespeichert. Sind alle Optionen mit dem Nutzerwissen verglichen, so werden die in der Liste gespeicherten Interaktionsoptionen mit höheren Wahrscheinlichkeiten belegt. Bevor nun durch einen Algorithmus eine Interaktionsoption ausgewählt werden kann, wird die Rule-Engine aufgerufen, die die Wahrscheinlichkeiten aller Interaktionsoptionen, basierend auf Regeln, verändert. Nach dem Aufruf der Rule-Engine werden die resultierenden Wahrscheinlichkeiten normalisiert. Ist die Normalisierung abgeschlossen wird durch einen Algorithmus entschieden welche Interaktionsoption ausgewählt wird. Die Auswahl erfolgt in Abhängigkeit der normalisierten Wahrscheinlichkeiten. Existieren beispielsweise nur zwei Interaktionsoptionen,

wovon die erste Option eine Wahrscheinlichkeit von $p = 0,99$ hat und die zweite Interaktionsoption eine Wahrscheinlichkeit von $p = 0,01$, so werden von 100 Iterationen im Mittel 99-mal die erste Option und ein einziges Mal die zweite Option gewählt.

Neben dem DPROM existieren noch drei weitere Module zur Verarbeitung der Informationen. Das Semantic-Processing-Module (SPM) leitet sich vom DPROM ab, es unterscheidet sich allerdings darin, wie die Informationen der Interaktionsoptionen mit dem Nutzerwissen verglichen werden. Wird das SPM in der Simulation verwendet, werden das Nutzermodellwissen und die modellierten Informationen durch Synonyme angereichert. Diese Anreicherung hat den Vorteil, dass das Aufgabenwissen von den modellierten Informationen der Benutzerschnittstelle abweichen kann. So kann z. B. ein Button ein Label mit dem Text *Fernseher* haben und das Nutzerwissen als Information zum zu verwendenden Gerät den Wert *Fernsehgerät* enthalten, ohne dass der Vergleich von Nutzerwissen und Interaktionsinformationen aufgrund dieses Unterschiedes fehlschlägt. Die Synonyme werden durch Anfragen an verschiedene Datenbanksysteme gefunden. Für die deutsche Sprache werden unter anderem die Datenbanken vom Webprojekt Wortschatz Leipzig³, OpenThesaurus⁴ und DISCO (extracting DIstributionally related words using CO-occurrences)⁵ verwendet. Es besteht weiterhin die Möglichkeit andere Dienste, wie GermaNet⁶, zu nutzen, wenn die entsprechenden Lizenzen vorliegen.

Das Multimodal Processing Module (MPM) ermöglicht die (semi-) automatische Evaluation von multimodalen Systemen. Multimodale Systeme sind über mehr als eine Modalität bedienbar. Das MPM kann, basierend auf einer Abschätzung der noch benötigten Interaktionsschritte entscheiden, ob die Eingabemodalität gewechselt wird, um eine Aufgabe mit einer anderen Modalität effizienter lösen zu können [205]. In der vorliegenden Implementierung kann das Nutzermodell zwischen Spracheingabe und Mauseingabe wechseln um eine Aktion auszulösen.

Das Utterance Processing Module (UPM) kann nur für Sprachdialogsysteme (SDS) verwendet werden. Das UPM nutzt zur Bestimmung der Anzahl der zu übermittelnden Konzepte (Informationen aus dem Nutzermodellwissen) ein statistisches Modell und verzichtet auf den Einsatz der Rule-Engine. Beim UPM wird basierend auf den Eigenschaften des Prompts entschieden, ob das Nutzermodell als Antwort einen vollständigen Satz, einen Teils eines Satzes oder ein einzelnes Wort an das System übermittelt.

Wurden vom gewählten Verarbeitungsmodul eine Interaktionsoption ausgewählt, wird diese Wahl an das Modul zur Ausführung der Interaktionsoption übergeben.

³ <http://wortschatz.uni-leipzig.de/>

⁴ <https://www.openthesaurus.de/>

⁵ <http://www.linguatools.de/disco/disco.html>

⁶ <http://www.sfs.uni-tuebingen.de/lsd/index.shtml>

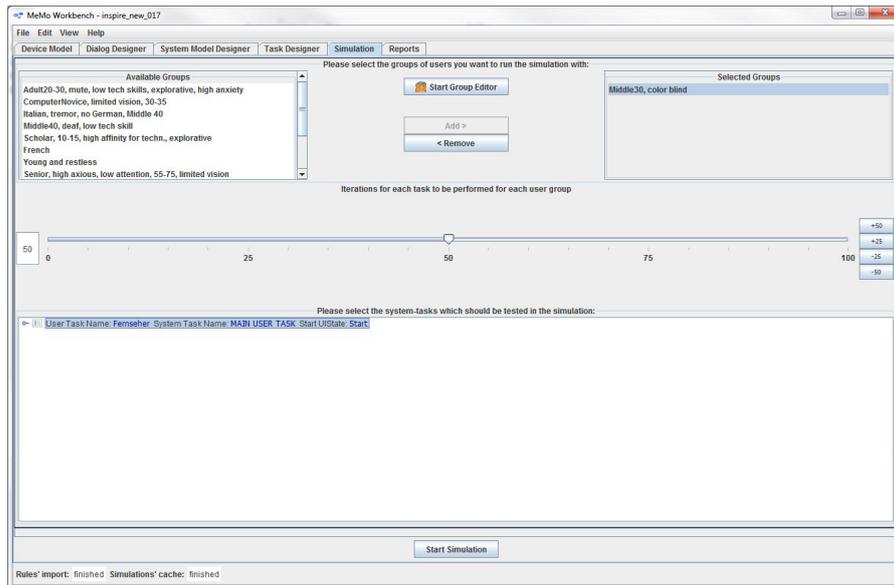


Abbildung 11: Fenster zur Festlegung der Nutzergruppen, Anzahl der Iterationen und der Aufgabenstellung

2.1.5.3 Ausführung

Das Modul zur Ausführung der berechneten Interaktionsoption wird dazu genutzt, die gewählte Interaktion auf das Eingabegerät anzuwenden. Hat sich das Nutzermodell für einen *Left click* entschieden, so wird diese Interaktion vom Ausführungsmodul auf das simulierte Eingabegerät angewendet. In der vorliegenden MeMo-Werkbank existiert nur ein Ausführungsmodul, das sogenannte Default Execution Module (DEM). Das DEM funktioniert, wie auch das DPERM, fehlerfrei. Es werden demnach keine Fehler durch falsche oder fehlerhafte Ausführungen von Interaktionen simuliert. Das Nutzermodell wird deshalb die Bedienelemente auf grafischen Benutzerschnittstellen immer treffen und bei SDS die zu übermittelnden Konzepte immer fehlerfrei aussprechen (senden).

2.2 SIMULATION

Die Simulation dient dazu, die Nutzermodelle mit dem Systemmodell interagieren zu lassen. Um die Simulation zu starten muss der Entwickler den in Abbildung 11 dargestellten Reiter in der Werkbank öffnen. Ist der Reiter *Simulation* geöffnet, werden oben auf der linken Seite die verfügbaren Nutzermodellgruppen angezeigt. In der rechten oberen Ecke des Reiters werden die vom Entwickler ausgewählten Nutzermodellgruppen aufgelistet. Die Nutzermodellgruppen können über einen Editor bearbeitet werden, der durch einen Button zwischen den beiden Listen aufgerufen werden kann. Unter den Nutzermodellgruppen ist ein Schieber, mit dem der Entwickler festlegen kann wie

viele Iterationen er durchführen will. Pro Iteration wird ein neues Nutzermodell aus den Nutzermodellgruppen abgeleitet. Hat der Entwickler 2 Gruppen für die Simulation ausgewählt und 50 Iterationen eingestellt, werden insgesamt 100 Iterationen simuliert, da die Angabe der Anzahl der Iterationen pro Nutzermodellgruppe gilt. Sind die Nutzermodellgruppen und die Anzahl der Iterationen festgelegt, so muss vom Entwickler noch bestimmt werden, welches Aufgabenmodell von den Nutzermodellen mit dem Systemmodell gelöst werden soll. Ist die Aufgabenstellung ausgewählt, kann die Simulation gestartet werden, indem der Button mit dem Label *Start Simulation* gedrückt wird.

2.2.1 Ablauf

Ist die Simulation vom Modellierer gestartet worden, wird von der MeMo-Werkbank die erste Nutzermodellgruppe aus der Liste der ausgewählten Gruppen verwendet, um das Nutzermodell für die erste Iteration zu generieren. Wie ein Nutzermodell aus der Beschreibung der Nutzermodellgruppe abgeleitet wird wurde im Abschnitt 2.1.5 beschrieben. Ist das Nutzermodell für die Iteration erzeugt, wird das Systemmodell mit den im Aufgabenmodell spezifizierten Informationen initialisiert und das Nutzermodell bekommt das Nutzerwissen (Informationen) für die erste Unteraufgabe übermittelt. Hiermit ist die Vorbereitung eines Iterationsdurchlaufs abgeschlossen.

Nach der Initialisierung der Simulation beginnt das Nutzermodell mit der Lösung der Aufgabe. Der erste Schritt in jeder Iteration ist die Wahrnehmung der möglichen Interaktionsoptionen in der Benutzerschnittstelle. Für die Wahrnehmung der Optionen können die in Abschnitt 2.1.5.1 beschriebenen Module verwendet werden, wobei die in der Simulation zu nutzenden Module vom Entwickler vor der Durchführung der Simulation festgelegt werden müssen. Wird hinsichtlich der Module keine Auswahl getroffen, so werden für die Wahrnehmung, Verarbeitung und Ausführung die Default Module verwendet.

Wurden die modellierten Informationen durch das entsprechenden Wahrnehmungsmodul erfasst, werden diese vom nachfolgenden Verarbeitungsmodul (siehe Abschnitt 2.1.5.2) ausgewertet. Während der Auswertung wird die Rule-Engine (siehe Abschnitt 2.1.2) eingesetzt, die die Werte für die Interaktionswahrscheinlichkeiten beeinflusst. Basierend auf der Auswertung der Benutzerschnittstelle und durch die Änderungen der Rule-Engine wird vom Verarbeitungsmodul eine Interaktionsoption ausgewählt und an das Modul zur Ausführung der Interaktion übergeben. Das vorliegende Modul zur Ausführung ist – wie in Abschnitt 2.1.5.3 beschrieben – relativ einfach aufgebaut und arbeitet fehlerfrei. Aufgrund der Fehlerfreiheit wird die vom Nutzermodell intendierte Interaktion direkt vom Ausführungsmodul auf das gewählte Bedienelement angewendet.

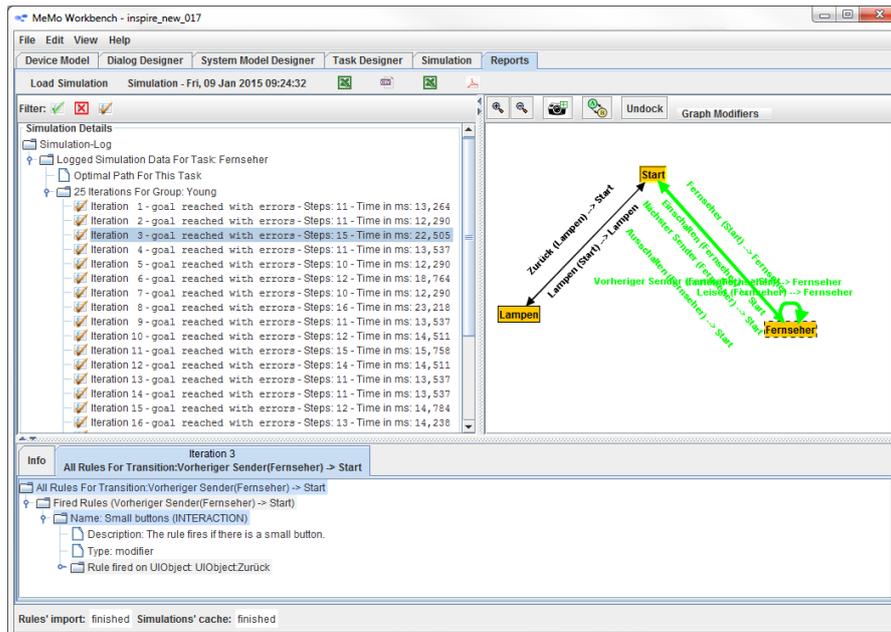


Abbildung 12: Report der MeMo-Werkbank für eine Iteration

Durch die Interaktion des Nutzermodells mit dem Systemmodell, wird die Verarbeitung im Systemmodell angestoßen. Es wird nun vom Systemmodell überprüft, ob eine, und wenn ja welche Transition ausgeführt werden kann. Wird eine Transition ausgeführt, so werden die Konsequenzen der Transition angewendet. Durch die Anwendung der Konsequenzen werden die Werte von Informationen verändert. Sind alle Konsequenzen verarbeitet, wird vom der MeMo-Werkbank überprüft, ob die Aufgabenstellung erfolgreich abgeschlossen werden konnte, das heißt, ob alle definierten Abbruchbedingungen erfüllt worden sind. Sind alle Abbruchbedingungen erfüllt, wird von der Werkbank geprüft, ob weitere Unteraufgaben existieren. Existieren weitere Aufgaben, so wird dem Nutzermodell das entsprechende Nutzerwissen übertragen und das Systemmodell wird, falls notwendig, mit neuen Startwerten initialisiert. Sind keine weiteren Unteraufgaben vorhanden, ist die aktuelle Iteration abgeschlossen und es wird eine neue Iteration gestartet, wenn dies erforderlich ist.

Während der Simulation wird in jedem Schritt aufgezeichnet, welche Regeln angewendet wurden und wie die Wahrscheinlichkeiten waren. Durch die Aufzeichnung der Daten kann der Entwickler nachvollziehen, wieso das Nutzermodell vom optimalen Pfad abgewichen ist und Rückschlüsse auf notwendige Veränderungen in der Gestaltung der Benutzerschnittstelle ziehen.

2.2.2 Ergebnisse

Ist die Simulation abgeschlossen, kann der Entwickler über den Klick auf den Reiter *Reports* die Ansicht mit den Simulationsergebnissen anzeigen lassen. In Abbildung 12 ist die Berichtsansicht dargestellt. Auf der linken Seite kann der Entwickler das Simulationsprotokoll öffnen und sich für jede Iteration anzeigen lassen, wie viele Interaktionsschritte vom Nutzermodell gebraucht wurden, ob die Lösung der Aufgabe fehlerfrei war, und wie lange die Ausführungszeit ist. Auf der rechten Seite wird der Interaktionspfad durch das Systemmodell visualisiert. Unter der Pfadvisualisierung und dem Simulationsprotokoll können Details zur Simulation angezeigt werden. Im dargestellten Beispiel sind die Details für die dritte Iteration beschrieben. Für diese Iteration ist dort zu lesen, dass bei der Aufgabe auf einen anderen Sender umzuschalten, eine Regel verarbeitet wurde, die aktiviert wird, wenn in der Benutzerschnittstelle ein Button vorhanden ist, der klein ist. Zusätzlich ist zu erkennen, dass die Regel aufgrund des *Zurück* Buttons ausgelöst wurde.

Neben der Darstellung der Simulationsergebnisse in der MeMo-Werkbank besteht die Möglichkeit eine Zusammenfassung als PDF zu drucken oder auch ein detaillierteres Protokoll als CSV-Datei auszugeben. Das PDF kann zur Dokumentation der Ergebnisse verwendet werden und die CSV-Datei für die Weiterverarbeitung in Statistikprogrammen.

2.3 FAZIT

In der vorliegenden Form ist die MeMo-Werkbank geeignet, um das *System Mental Model* (vgl. Abschnitt 1.3.3) zu evaluieren.

Mit der vorliegenden Werkbank können keine Fehler im *Device Mental Model* simuliert werden, da in der aktuellen Version nur Systemmodelle evaluiert werden können, die per Sprache oder Maus bedient werden. Weiterhin ist die Abschätzung der Schwere eines gefundenen Usability-Problems schwierig, da die verwendeten Nutzermodelle keine demografischen Verteilungen widerspiegeln. In gleicher Weise ist es für den Entwickler schwierig, die Zusammenhänge zwischen einzelnen Einschränkungen eines Nutzers herzustellen. Speziell bei älteren Menschen mit Multimorbidität – dem gleichzeitigen Vorliegen von mehreren Erkrankungen [206] – ist die Definition von demografisch validen Nutzermodellen ein großes Problem.

Um die genannten Probleme zu lösen muss die MeMo-Werkbank, um mehrere Aspekte, erweitert werden. Zum einen muss die Definition und Generierung von demografisch validen Nutzermodellen vereinfacht werden, ohne den Entwickler zu stark in seinen Freiheiten einzuschränken. Zum anderen müssen mit der MeMo-Werkbank auch Fehler aufgrund von falschen oder fehlenden mentalen Modellen be-

zöglich der Bedienung von Eingabegeräten, die sogenannten *Device Mental Model*, simuliert werden können.

Das Ziel dieses Kapitels ist die Dokumentation der Lösung, wie vorhandene Daten über Einschränkungen in der Bevölkerung für die automatische Evaluation der Usability genutzt werden können. Der erste Abschnitt in diesem Kapitel beschreibt die Nutzerattribute, die in die technische Lösung integriert werden. Zusätzlich wird beschrieben, welche Ausprägungen die einzelnen Einschränkungen haben, wie sich diese Einschränkungen auf die Interaktion auswirken, und wie hoch die jeweiligen Anteile der Ausprägungen in der Bevölkerung sind. Der zweite Abschnitt in diesem Kapitel stellt die technische Lösung – das Bayes'sche Netz – vor. In Abschnitt 3.2.1 wird zuerst das Konzept der Bayes'schen Netze vorgestellt und anhand von Beispielen erläutert. Daran anschließend wird der Aufbau der konkreten Lösung und die Umsetzung als Webservice vorgestellt. Den Abschluss dieses Kapitels bildet die Validierung der technischen Lösung.

3.1 NUTZERATTRIBUTE

Wie in Abschnitt 1.4.2 dargelegt können Nutzermodelle unterschiedliche Ausprägungen haben und verschiedene Attribute integrieren. Damit ein Nutzermodell einen realen Nutzer repräsentieren kann muss es die Ausprägungen der Eigenschaften eines Nutzers abdecken können.

Nach Kujala und Kauppingen [142] müssen für die Auswahl von Teilnehmern an einem empirischen Versuch verschiedene Nutzerattribute betrachtet werden, damit der Versuch die Heterogenität der Bevölkerung abdecken kann und valide Ergebnisse erzielt werden können. Die Autoren nennen in ihrer Arbeit drei Kategorien anhand derer die Versuchspersonen ausgewählt werden sollen, um möglichst repräsentativ zu sein. Die erste Kategorie, die *persönlichen Merkmale*, umschreibt Attribute wie Geschlecht, Alter, Rolle in der Organisation, Persönlichkeit, Fähigkeiten, Einschränkungen und den sozioökonomischen Status. Die Ziele und Motivation einer Versuchsperson, die Aufgabenstellung und die Vorerfahrung gliedern die Autoren in die Kategorie der *aufgabenorientierten Eigenschaften einer Person*. In die dritte Kategorie, die *geografischen und sozialen Eigenschaften einer Person*, fallen Angaben über die geografische Herkunft einer Versuchsperson, deren kulturelle Hintergründe und ihre soziale Umgebung.

In dieser Arbeit werden vor allem Merkmale aus der ersten Kategorie, den persönlichen Merkmalen, genutzt, um Nutzermodelle zu beschreiben. Zusätzlich werden größtenteils nur die Merkmale

betrachtet, die die Interaktion mit einem System negativ beeinflussen können. Die Begründung für die Fokussierung auf diesen Bereich motiviert sich auf dem Ziel der Arbeit, das Interaktionsverhalten von älteren Personen während der Nutzung von mobilen Endgeräten zu simulieren.

Nach Persad et al. [178] beeinflussen besonders Einschränkungen im Hören, im Sehen, in der Kognition und in der Motorik das Verhalten von Versuchspersonen während der Interaktion, es existieren allerdings auch weitere Einschränkungen, die keiner der genannten Kategorien zugehören. In den nachfolgenden Abschnitten der Arbeit wird zu jeder Einschränkungskategorie dargestellt wie sie sich auswirkt, wie die Verteilung in der Bevölkerung ist, und welche Zusammenhänge mit anderen Attributen bestehen. Das Ziel dieses Teils der Arbeit ist es, einen Überblick über die Zusammenhänge und Heterogenität der möglichen Einschränkungen zu geben.

3.1.1 Hören

Neben der visuellen Wahrnehmung ist das Hören der wichtigste Sinn, wenn es darum geht mit einem System zu interagieren. Gleichwohl ist der Gehörsinn bei vielen Personen beeinträchtigt, was die Interaktion erschweren kann.

In diesem Abschnitt wird ein Überblick über das Thema Schwerhörigkeit gegeben. Es werden Faktoren genannt, die das Risiko für eine Schwerhörigkeit erhöhen und es wird beschrieben, wie sich ein Hörverlust auf die Interaktion auswirken kann. Der Abschnitt schließt mit der Darstellung der Daten zur Verteilung von Schwerhörigkeit in der Gesellschaft.

3.1.1.1 Überblick

Damit etwas gehört werden kann, muss durch ein Ereignis ein Schall erzeugt werden [7, S. 120 ff]. Der erzeugte Schall trifft auf das Ohr und wird dort, in der Hörschnecke (Cochlea), zum Hörereignis indem der Schall in Nervenimpulse umgewandelt wird. Von der Cochlea werden die Nervenimpulse über den Hörnerv in die auditive Rinde weitergeleitet, wo die Signale weiterverarbeitet werden.

In der Literatur wird zwischen *leichter*, *mittlerer* und *schwerer Schwerhörigkeit* unterschieden [35, 98]. Nach Hesse und Laubert [98] bedeutet eine geringe Schwerhörigkeit einen Hörverlust von 20 bis 40 Prozent, eine mittelschwere Schwerhörigkeit liegt bei einem Hörverlust von 40 bis 60 Prozent vor und als hochgradig schwerhörig gilt wer 60 bis 80 Prozent seiner Hörfähigkeit verloren hat. Wird die Schwerhörigkeit anhand der Hörschwelle eingeteilt, so werden Hörverluste von unter 40 Dezibel (dB) als *mild*, zwischen 41 und 55 dB als *moderat*, von 56 bis 70 dB als *moderat-ausgeprägt* und zwischen 70 und 90 dB als *ausgeprägt* beschrieben werden [250].

Bei den Einschränkungen im Hören sind insbesondere die altersbedingten Einschränkungen ein weit verbreitetes Phänomen. Nach Walling und Dickson [242] bzw. Wascher et al. [244] sind rund ein Drittel der älteren Erwachsenen zwischen 61 und 70 Jahren, knapp 40 % der Personen über 70 und über 80 % der Personen älter als 85 Jahre von Höreinschränkungen betroffen.

3.1.1.2 *Risikofaktoren und Ursachen für Einschränkungen*

Als Risikofaktoren gelten nach Chou et al. [35] eine häufige Belastung durch Lärm, genetische Faktoren, die Einnahme von gehörschädigenden Medikamenten, häufige Innenohrinfektionen, Diabetes mellitus und Bluthochdruck. Zusätzlich zeigt sich, dass Männer häufiger als Frauen, und Personen aus sozial schwache Schichten stärker als Personen aus anderen Schichten, von Schwerhörigkeit betroffen sind. Eine hohe und ständige Lärmbelastung ist die häufigste Ursache für eine Hörstörung [226].

Ein Großteil der Schwerhörigkeit im Alter (Presbyakusis) ist auf die Degeneration von Haarzellen im Innenohr und auf Abbauprozesse der zentralen Hörbahnen zurückzuführen [65, S. 352]. Allerdings können auch andere Ursachen (z. B. beim neuronalen Typ, durch Degeneration von Ganglienzellen) eine Schwerhörigkeit bedingen [208]. Die Presbyakusis zeichnet sich dadurch aus, dass sie beidseitig und symmetrisch ist [35]. Die Hörverluste sind über alle Frequenzen zu beobachten, allerdings sind höhere Frequenzen stärker von der Degeneration betroffen [98].

3.1.1.3 *Auswirkungen*

Eine Einschränkung im Hören hat direkte und indirekte Auswirkungen auf die betroffene Person. Zu den direkten Auswirkungen zählen u. a. Schwierigkeiten beim Sprachverstehen und Probleme in der Lokalisierung einer Schallquelle [35]. Nach Persad et al. [178] sind es insbesondere diese Einschränkungen, die die Interaktion mit einem System stark beeinträchtigen. Indirekt kann eine Schwerhörigkeit zur Zunahme bei emotionalen Dysfunktionen bzw. zu Depressionen und zur sozialen Isolation führen [35]. Emotionen haben, nach Brave und Nass [21], einen Einfluss darauf wie Menschen mit einem System interagieren. Darüber hinaus beeinflussen Emotionen den Lernprozess. Negative Gefühle während des Lernens führen eher dazu, dass negative Aspekte der Interaktion gemerkt werden

3.1.1.4 *Verteilungen von Einschränkungen in der Bevölkerung*

Nach Agrawal et al. [1] liegt der Anteil der Personen, die eine Einschränkung im Hören haben, in den USA bei etwa 16,1 %. In Deutschland wird davon ausgegangen, dass etwa 17 % der Bevölkerung von einer Höreinschränkung betroffen sind [95, S. 70]. Wie in

Altersgruppe in Jahren	Anteil Schwerhöriger			
	Deutschland	England ^a	Italien ^b	USA
14 – 19	1 %	-	-	-
17 – 30 ^a				
18 – 30 ^b				
20 – 29	2 %	1,8 %	1,9 %	3,1 %
30 – 39				
31 – 40 ^{a,b}	5 %	2,8 %	3,9 %	5,4 %
40 – 49				
41 – 50 ^{a,b}	6 %	8,2 %	8,3 %	15 %
50 – 59				
51 – 60 ^{a,b}	25 %	18,9 %	18,7 %	29 %
60 – 69				
61 – 70 ^{a,b}	37 %	36,8 %	37,7 %	49 %
70+				
71+ ^{a,b}	54 %	60,2 %	69,4 %	-

Tabelle 4: Verteilung der Schwerhörigkeit nach Altersgruppen

den vorangegangenen Unterabschnitten beschrieben, sind ältere Menschen stärker von Hörverlusten betroffen als jüngere Menschen.

In Tabelle 4 ist die Verteilung der Schwerhörigkeit in verschiedenen Ländern dargestellt. Die Zahlen für Deutschland stammen aus einer Studie mit 2031 Teilnehmern, die von Sohn und Jörgenshaus im Jahr 2001 durchgeführt wurde [217]. Die Teilnehmer der deutschen Studie wurden in 11 Allgemeinarztpraxen untersucht. Als schwerhörig wurde eingestuft, wer bei Frequenzen von 500, 1000, 2000, 3000 und 4000 Hz und einem Schalldruckpegel von 40 dB, einen der Töne nicht wahrnehmen konnte.

Die Daten für England stammen aus der Studie von Davis [47] und basieren auf 2662 Audiogrammen. In dieser Studie wurde ein Teilnehmer als schwerhörig klassifiziert, wenn der Hörverlust auf dem besseren Ohr mindestens 25 dB betrug. Das Mindestalter der Studienteilnehmer wurde mit 17 Jahren definiert, weshalb die erste Altersgruppe die Spanne von 17 bis 30 Jahre umfasst. Wie in der Tabelle dargestellt, weicht die Definition der Altersklassen, um ein Jahr von der deutschen Studie ab.

Für die italienischen Daten wurde auf die Studie von Quaranta et al. [182] aus dem Jahr 1995 zurückgegriffen, bei der die Grenze zur Schwerhörigkeit ebenfalls bei 25 dB lag. Das Mindestalter für die

Teilnahme an der Studie lag bei 18 Jahren, weshalb die erste Altersgruppe die Teilnehmer von 18 bis 30 Jahren enthält. Die Definition der Altersgruppe entspricht der Studie von Davis. Für die Studie von Quaranta et al. wurden insgesamt 2170 Probanden untersucht.

Die Daten zur Verteilung in den USA stammen aus der Studie von Agrawal et al. [1]. Agrawal et al. haben ebenfalls einen Hörverlust von mindestens 25 dB, auf dem besseren Ohr, in den für das Sprachverstehen wichtigen Frequenzen, als Kriterium für die Bewertung als Schwerhöriger angenommen. Insgesamt wurden 5742 Personen untersucht. Nach Agrawal et al. beträgt das Quotenverhältnis zwischen Männern und Frauen, über alle Altersgruppen hinweg, 2,4 : 1. Das heißt, dass Männer mehr als doppelt so häufig von Schwerhörigkeit betroffen sind, wie es Frauen sind.

Neben dem Geschlecht führen auch andere Erkrankungen, wie Diabetes mellitus, zu einem höheren Prozentsatz an Schwerhörigen. In der Studie von Bainbridge et al. [8] konnte gezeigt werden, dass der Anteil der Personen mit einer Höreinschränkung, in der Gruppe mit Diabetikern nahezu doppelt so hoch ist, wie in der Vergleichsgruppe, in der die Personen keinen Diabetes mellitus haben.

Weiterhin kann eine ungesunde Lebensweise zu Höreinschränkungen führen. In der Studie von Cruickshanks et al. [42] wurde gezeigt, dass der Anteil von Personen mit Höreinschränkungen unter aktiven Rauchern etwa 2-mal höher (Faktor 1,69) liegt, als unter Nichtrauchern.

Als letzte Faktoren sollen der Bildungsstand und das Einkommen betrachtet werden. In einer weiteren Studie von Cruickshanks et al. [43] zeigte sich, dass für die Gruppe der Personen ohne High School Abschluss die Wahrscheinlichkeit, dass eine Höreinschränkung vorliegt, um den Faktor 2,42 höher ist. Ähnliche Ergebnisse liefert das Jahreseinkommen. Hier zeigt sich, dass Personen die weniger als 30.000,- Dollar pro Jahr verdienen, etwa doppelt so oft eine Höreinschränkung haben, wie Personen die über 60.000,- Dollar pro Jahr verdienen [43].

3.1.1.5 Zusammenfassung

In diesem Unterabschnitt wurde gezeigt, dass die Verteilung für Höreinschränkungen vom Alter, dem Geschlecht, der Lebensweise, dem Bildungsniveau und anderen Erkrankungen (z. B. Diabetes mellitus) beeinflusst wird. Weiterhin konnte gezeigt werden, dass Einschränkungen im Hören kein seltenes Problem sind, sondern dass nahezu jeder 5. Mensch hiervon betroffen ist. Aufgrund der demografischen Entwicklung ist davon auszugehen, dass der Anteil von Personen mit Einschränkungen beim Hören weiter steigen wird.

3.1.2 *Sehen*

Die visuelle Wahrnehmung ist der wichtigste Sinn, um mit einem interaktiven System zu interagieren [13, 178]. In diesem Unterabschnitt werden die wichtigsten Seheinschränkungen vorgestellt und deren Auswirkungen auf die Interaktion mit einem System beschrieben.

3.1.2.1 *Überblick*

Die Auswertung von visuellen Informationen beginnt mit der Verarbeitung von Licht an der inneren Oberfläche des Auges, der sogenannten Retina [13]. Das durch das Auge wahrgenommene Licht wird in neuronale Signale umgewandelt, welche in das Gehirn weitergeleitet werden. Im Gehirn erreichen die neuronalen Signale zuerst die primäre Sehrinde, welche die wahrgenommenen Informationen abbildet und einfache visuelle Analysen ermöglicht (Kantenerkennung, Ausrichtung und Farbunterschiede). Die weitere Verarbeitung der wahrgenommenen visuellen Informationen erfolgt in den angrenzenden Bereichen.

Nach Lange [148] können verschiedene angeborene oder erworbene Beeinträchtigungen zu Einschränkungen im Sehen führen. Die häufigsten Beeinträchtigungen sind Fehlsichtigkeit (Kurzsichtigkeit, Weitsichtigkeit), Grauer Star (Katarakt, meist altersbedingt Trübung der Linse), verschiedenen Netzhauterkrankungen (z. B. altersabhängige Makuladegeneration) und der Grüne Star (Glaukom). Darüber hinaus können Wahrnehmungsstörungen nach Verletzungen im Gehirn auftreten (z. B. nach einem Schlaganfall [13]).

3.1.2.2 *Risikofaktoren und Ursachen für Einschränkungen*

Eine Einschränkung in der visuellen Wahrnehmung kann angeboren oder erworben sein [148]. Ein Beispiel für eine angeborene Einschränkung ist die einfache Myopie (Kurzsichtigkeit) [83, S. 336]. Erworbene Einschränkungen können die Folge von neurologischen oder allgemeinen Erkrankungen (z. B. Diabetes mellitus) sein [148].

In verschiedenen Studien [148, 231] wurde gezeigt, dass mit zunehmenden Alter der Anteil der Personen mit Sehbeeinträchtigungen zunimmt. So treten Augenerkrankungen, wie der Grauer Star (Katarakt) oder die altersabhängige Makuladegeneration (AMD), signifikant häufiger bei hochaltrigen Menschen (80 Jahre und älter) auf [39, 174].

Neben diesem Alterseffekt, existieren auch Unterschiede zwischen den Geschlechtern. In einigen Studien [38, 135, 148] konnte gezeigt werden, dass Frauen ein höheres Risiko für eine Seheinschränkung haben. Nur bei der diabetischen Retinopathie ist das Risiko für Männer höher als das der Frauen [258].

Wie schon bei den Hörbeeinträchtigungen in Abschnitt 3.1.1 zeigt sich auch bei den Einschränkungen in der visuellen Wahrnehmung, dass Raucher stärker von Beeinträchtigungen betroffen sind. So stellten

Thornton et al. [233] fest, dass bei der altersabhängigen Makuladegeneration (AMD), das Risiko für Raucher höher als für Nichtraucher ist.

3.1.2.3 Auswirkungen

Einschränkungen in der visuellen Wahrnehmung beeinträchtigen meistens die Interaktion mit einem System (direkte Auswirkungen). Neben den direkten Auswirkungen von Sehbeeinträchtigungen existieren auch indirekte Auswirkungen. Zu den indirekten Auswirkungen zählen Einschränkungen im physischen, emotionalen und sozialen Wohlbefinden [148, S. 63]. Nachfolgend soll allerdings nur auf die direkten Auswirkungen eingegangen werden.

Bei einer Kurzsichtigkeit (Myopie) können visuelle Reize im Nahbereich meist problemlos wahrgenommen werden, während weiter entfernte Reize nur unscharf zu erkennen sind. Demgegenüber steht die Weitsichtigkeit (Hyperopie), bei der eine Person ein Objekt in der Ferne gut erkennen kann, im Nahbereich allerdings auf eine Sehhilfe angewiesen ist, um verschiedene visuelle Reize unterscheiden zu können [63, S. 383]. Die Einschränkungen, die durch eine Myopie und Hyperopie entstehen, können meist durch Sehhilfen ausgeglichen werden, allerdings ist es problematisch, wenn der Visus zwischen nahen und fernen Interaktionsobjekten wechseln muss und die Akkommodationsfähigkeit des Auges (Flexibilität der Linse) nicht ausreicht.

Bei einem Grauen Star trübt sich die Augenlinse ein. Durch die Eintrübung werden visuelle Reize nur noch verschwommen wahrgenommen, die Sehschärfe nimmt ab [63, S. 382]. Kleinere visuelle Abweichungen können nicht wahrgenommen werden, was die Unterscheidbarkeit von Objekten erschwert oder verhindert. Weiterhin kommt es bei einem Grauen Star vermehrt zu einer erhöhten Blendempfindlichkeit und zu einer blässeren Farbwahrnehmung.

Die altersabhängige Makuladegeneration (AMD) führt, durch die Schädigung der Makula, häufig zum Verlust der Lesefähigkeit [133]. Da die AMD den Bereich des schärfsten Sehens betrifft können auch Aufgaben nicht gelöst werden, die eine Unterscheidung von kleineren bis mittleren Details erfordern.

Ein Glaukom schädigt den Sehnerv und führt zu Verlusten im Gesichtsfeld, die bis hin zur Erblindung führen können [63, S. 381]. Fallen die Verluste in frühen Phasen häufig nicht auf, so schränken die Gesichtsfeldverluste in späteren Phasen die Wahrnehmung massiv ein. So können visuelle Reize im peripheren Sichtfeld von Menschen ohne Glaukom normalerweise gut wahrgenommen werden; bei Personen mit Glaukom hängt die Wahrnehmung davon ab, ob der Bereich vom Gesichtsfeldverlust betroffen ist oder nicht.

Altersgruppe in Jahren	Sehbeeinträchtigungen			
	keine	leicht	schwerwiegend	blind
18 – 29	87,6 %	10,8 %	0,5 %	1,1 %
30 – 44	85,6 %	12,5 %	0,8 %	1,0 %
45 – 64	75,5 %	20,9 %	2,3 %	1,3 %
65+	66,0 %	26,0 %	4,7 %	3,3 %

Tabelle 5: Verteilung der Sehbeeinträchtigungen nach Altersgruppen in Deutschland für Frauen [148]

3.1.2.4 Verteilungen von Einschränkungen in der Bevölkerung

Für die Übersichtsstudie von Lange [148] wurden etwa 25.000 Menschen in Deutschland per Telefon hinsichtlich der allgemeinen gesundheitlichen Lage interviewt. Im Verlauf des Telefoninterviews gaben rund 16,5 % an, dass sie leichte Schwierigkeiten beim Sehen haben. Weiterhin gaben rund 2,3 % der Befragten an, dass sie große Schwierigkeiten beim Sehen haben. Rund 1,2 % der Befragten gaben an, dass sie nicht mehr sehen können. In Tabelle 5 sind die Daten für Frauen in verschiedene Altersgruppen dargestellt. Die gleichen Daten sind in Tabelle 6 für die Männer verzeichnet. Es ist festzustellen, dass Frauen häufiger von Sehschwierigkeiten berichten als Männer. Von leichten Sehschwierigkeiten berichten 18,5% der Frauen und 14,3% der Männer. Große Sehschwierigkeiten werden von 2,3% der Frauen und 1,2% der Männer angegeben. Einen vollständigen Verlust der Sehfähigkeit gaben 1,7% der Frauen und 0,7% der Männer an. Zusätzlich zum Unterschied zwischen den Geschlechtern kann eine Zunahme ab dem 45. Lebensjahr bei den Sehschwierigkeiten beobachtet werden. Ab dem 65. Lebensjahr nimmt der Anteil der Personen mit einem totalen Sehverlust stark zu.

In der Studie von Taylor et al. [231] wurde die Verteilung von Seheinschränkung und deren Gründe untersucht. Die Ergebnisse der Studie sind in Tabelle 7 aufgeführt. Nach Taylor et al. hat eine Person eine Sehschwäche, wenn, trotz Korrektur durch eine Sehhilfe, die Sehschärfe unter 50 % liegt. Liegt die Sehschärfe, trotz Sehhilfe, unter 10 % so wurde der Proband als blind bzw. nahezu blind eingestuft. Wie in der Tabelle dargestellt, ist in 62 % der Fälle der Auslöser für eine Sehbeeinträchtigung ein unkorrigierter *Refraktionsfehler*. Unter der Bezeichnung *Refraktionsfehler* sind Sehbeeinträchtigungen zu verstehen, die durch die Brechung des Lichts im Auge entstehen. Die meisten Refraktionsfehler lassen sich durch geeignete Sehhilfen (Brille) beheben [83]. Nach Grehn [83] sind rund 30 % der Menschen leicht kurzsichtig (weniger als -6 Dioptrien) und rund 2,5 % stark kurzsichtig (-6 Dioptrien und mehr).

Altersgruppe in Jahren	Sehbeeinträchtigungen			
	keine	leicht	schwerwiegend	blind
18 – 29	90,1 %	8,9 %	0,5 %	0,5 %
30 – 44	89,3 %	9,4 %	0,6 %	0,6 %
45 – 64	80,2 %	17,5 %	1,9 %	0,4 %
65+	76,9 %	20,6 %	1,3 %	1,2 %

Tabelle 6: Verteilung der Sehbeeinträchtigungen nach Altersgruppen in Deutschland für Männer [148]

Gründe	Auswirkung	
	Sehschwäche	Blind
Altersbedingte Makuladegeneration	10 %	48 %
Glaukom	3 %	14 %
Katarakt	14 %	12 %
Diabetische Retinopathie	2 %	11 %
Andere Netzhauterkrankungen	3 %	
Neurologisch	2 %	3 %
Refraktionsfehler	62 %	4 %
Andere	4 %	9 %

Tabelle 7: Gründe für Sehbeeinträchtigungen [231]

In der Meta-Studie der Eye Diseases Prevalence Research Group [129] zeigte sich, dass 40,3 % der Personen mit Diabetes mellitus eine diabetische Retinopathie haben. Bei 8,2 % der Personen wurde durch den Diabetes eine Retinopathie hervorgerufen, die zu Einschränkungen in der Wahrnehmung führte. Signifikante Unterschiede zwischen den Geschlechtern hinsichtlich der Auswirkungen von Diabetes konnten in dieser Studie nicht gefunden werden. In einer neueren Studie von Zhang et al. [258] sind die Anteile der Erkrankten deutlich geringer. Nach Zhang et al. haben 28,5 % der Menschen mit Diabetes mellitus eine diabetische Retinopathie und 4,4 % eine Einschränkung in der Wahrnehmung durch die Erkrankung. Weiterhin haben laut Zhang et al. Männer ein signifikant höheres Risiko, eine diabetische Retinopathie zu entwickeln, als dies Frauen haben. Die Unterschiede zwischen den Studien resultieren, mit hoher Wahrscheinlichkeit, aus den unterschiedlichen Kriterien für die Studienteilnehmer. Während in der Meta-Studie der Eye Diseases Prevalence Research Group [129] nur Personen ab 40 Jahren betrachtet wurden, fassen Zhang et al. [258] die Studie breiter und betrachten auch Teilnehmer ab 20 Jahren.

3.1.2.5 Zusammenfassung

In diesem Unterabschnitt wurde gezeigt, dass Einschränkungen in der visuellen Wahrnehmung zum Großteil auf unkorrigierten Refraktionsfehlern beruhen. Weiterhin wurde gezeigt, dass mit zunehmendem Alter der Anteil der Personen mit Sehbeeinträchtigungen steigt. Neben der erblichen Veranlagung zählen allgemeine Erkrankungen zu den häufigsten Gründen, die zu einer Seheinschränkung führen können.

3.1.3 Kognition

Der Begriff *Kognition* beschreibt die intelligente Verarbeitung der wahrgenommenen Informationen [169]. Zur Verarbeitung von Informationen werden, in Analogie zum Computer, ein Speicher (Gedächtnis) und ein Prozessor benötigt. Beide Teile können in ihrer Funktionsfähigkeit limitiert sein. In diesem Abschnitt werden verschiedene Einschränkungen und deren Auswirkungen auf das Interaktionsverhalten dargestellt.

3.1.3.1 Überblick

Bevor Informationen/Umweltreize verarbeitet werden können, müssen diese über die Sinne wahrgenommen worden sein. Im Umkehrschluss bedeutet das, dass ein Reiz nicht verarbeitet werden kann, wenn das entsprechende Sinnesorgan, z. B. durch eine Einschränkung, nicht in der Lage ist den Reiz wahrzunehmen. Wird ein Reiz wahrgenommen, werden die Reizinformationen durch die Sinnesorgane in Muster kodiert, die vom Gehirn verarbeitet werden. Wird die Analogie

des Computers zur Beschreibung der Kognition genutzt, so erfolgt die Verarbeitung der Muster mithilfe von *Prozessen* und durch Abrufe aus dem *Speicher*. Beide Komponenten können in ihrer Funktionsfähigkeit limitiert sein.

Eine Limitierung in den *Prozessen* betrifft die fluide Intelligenz. Der Begriff *fluide Intelligenz* beschreibt die Fähigkeit die wahrgenommenen Informationen schnell und akkurat im Kurzzeitgedächtnis zu repräsentieren und zu verarbeiten [73, S. 37]. Die Anlagen für die fluide Intelligenz werden vererbt [71, S. 50], können allerdings durch Training verbessert werden [111]. Das Problem mit der fluiden Intelligenz ist, dass die Leistungsfähigkeit nicht unbegrenzt ist und mit zunehmenden Alter abnimmt [23, 105]. Einschränkungen in diesem Bereich führen dazu, dass Informationen langsamer und weniger akkurat verarbeitet werden, was zu langsameren Ausführungszeiten und mehr Fehlern führen kann. Weiterhin verringert sich die Lernfähigkeit mit der abnehmenden Leistungsfähigkeit.

Unter dem Begriff *Speicher* ist die kristalline Intelligenz zu verstehen. Anders als die fluide Intelligenz basiert die kristalline Intelligenz auf durch Lernen erworbenem Wissen und ist somit sozial geprägt [71, S. 50]. Weiterhin ist die kristalline Intelligenz nicht so stark von den altersbedingten Einschränkungen betroffen und im Allgemeinen bei den älteren Menschen höher als bei den jüngeren Menschen [105, 200]. Die Einschränkungen in der kristallinen Intelligenz resultieren meist aus Erkrankungen (Demenz, Alzheimer), Lernproblemen oder Bildungsdefiziten.

3.1.3.2 Risikofaktoren und Ursachen für Einschränkungen

Kognitive Einschränkungen können durch eine Reihe von Faktoren bedingt sein. Da nicht alle an dieser Stelle aufgeführt werden können, werden nur die häufigsten Erkrankungen und deren Ursachen beschrieben.

Wie bereits in Abschnitt 3.1.3.1 dargestellt, nimmt die Leistungsfähigkeit der fluiden Intelligenz mit dem Alter ab. Hierbei handelt es sich oftmals nicht um Veränderungen aufgrund von Krankheiten oder Unfällen. Vielmehr summieren sich die über das Lebensalter zunehmenden kognitiven Einschränkungen, so dass im höheren Alter stärkere Beeinträchtigungen auftreten [200].

Unter dem Begriff Demenz ist eine Erkrankung des Gehirns zu verstehen, bei der die Hirnfunktionen eingeschränkt bzw. gestört sind. Nach der Definition der WHO ist „Demenz [...] ein Syndrom als Folge einer meist chronischen oder fortschreitenden Krankheit des Gehirns mit Störung vieler höherer kortikaler Funktionen, einschließlich Gedächtnis, Denken, Orientierung, Auffassung, Rechnen, Lernfähigkeit, Sprache und Urteilsvermögen“ [255]. Als Risikofaktoren für eine Demenz gelten nach Maercker [159, S. 60]: eine geringe Schulbildung, wenig körperliche und geistige Aktivität, Rauchen, Adipositas,

übermäßiger Alkoholkonsum, Stress, psychische Erkrankungen, Bluthochdruck, Diabetes mellitus, genetische Veranlagungen, Diäten und Vitamin-B12- und Folsäuremangel.

Neben der Demenz ist die Intelligenzminderung – die unvollständige geistige Entwicklung, die durch einen Intelligenzquotienten (IQ) von unter 70 definiert ist – eine häufig zu beobachtende kognitive Einschränkung. Die Ursache für eine Intelligenzminderung ist, abhängig vom Schweregrad, in 30 bis 60 % der Fälle unbekannt [72, S. 223]. Männer sind, in der Gruppe mit leichten Intelligenzminderungen, stärker vertreten als Frauen, ebenso ist der Anteil der Personen mit einer geistigen Behinderung in den niedrigen sozialen Schichten höher. Die Ursachen für schwere und schwerste Grade von geistiger Behinderung liegen meist vor der Geburt, allerdings kann hier in 30 % der Fälle die Ursache nicht bestimmt werden [243, S. 508].

Neben den bereits genannten Ursachen, entstehen kognitive Einschränkungen auch durch Unfälle oder Schlaganfälle. In diesem Zusammenhang wird von *erworbenen Hirnschäden* und nicht von geistiger Behinderung gesprochen [212, S. 16]. In Deutschland haben pro Jahr etwa 550.000 Menschen einen Unfall/eine Erkrankung, die zu Hirnschäden führt [40, S. 9]. Die häufigsten Erkrankungen in diesem Zusammenhang sind der Schlaganfall und das Schädel-Hirn-Trauma. Männer haben ein höheres Risiko einen Schlaganfall zu bekommen, allerdings haben absolut gesehen, aufgrund der demografischen Verteilung, mehr Frauen als Männer einen Schlaganfall [187]. Neben dem Geschlecht wird auch die Zugehörigkeit zu einer niedrigen sozialen Schicht als Risikofaktor genannt [26]. Als weitere beeinflussbare Risikofaktoren gelten Bluthochdruck, Herzrhythmusstörungen, Rauchen, Diabetes mellitus, eine hoher Cholesterinspiegel, übermäßiger Alkoholkonsum und wenig körperliche Aktivität [179, S. 5]. Das Risiko ein Schädel-Hirn-Trauma (SHT) zu erleiden ist im Alter von 15 bis 25 und bei älteren Menschen über 65 Jahren besonders hoch, darüber hinaus sind Männer doppelt so häufig betroffen wie Frauen [153, S. 579].

3.1.3.3 Auswirkungen

Kognitive Einschränkungen beeinflussen die Interaktion mit einem System auf vielfältige Weise. Durch Einschränkungen im Gedächtnis werden Informationen nicht oder nur unzureichend erinnert, was zu Problemen in der Interaktion führen kann. Ist das Denken oder die Auffassungsgabe beeinträchtigt, so fällt es der Person schwer Zusammenhänge zu erkennen bzw. notwendige Schritte abzuleiten. Bei Einschränkungen in der Lernfähigkeit dauert es länger bis eine Person die Bedienung des Systems verinnerlicht hat; häufig sind sich wiederholende Fehler zu beobachten.

Eine Demenzerkrankung beeinträchtigt die Orientierung, Aufmerksamkeit und Sprache der betroffenen Person. Für die Interaktion mit einem System besonders relevant und durch eine Demenz stark beein-

trächtig sind insbesondere das Gedächtnis (Kurzzeitgedächtnis und Arbeitsgedächtnis) und die exekutiven Funktionen (Planen/Problemlösen und kognitive Flexibilität) [210, S. 238].

Nach einem Schlaganfall haben rund 25 % der Patienten schwere Einschränkungen [99]. Abhängig von dem geschädigten Gebiet, können Sprachstörungen, Sprachverständlichkeitsstörungen, Bewegungsstörungen, Gesichtsfeldausfälle, Gedächtnisstörungen und Probleme bei der Bewegungskoordination auftreten [179, S. 7].

Zu den Beeinträchtigungen nach einem Schädel-Hirn-Trauma zählen Gedächtnisprobleme, schnellere Ermüdung, leichte Irritierbarkeit, Langsamkeit und Konzentrationsprobleme [170, S. 112]. Weiterhin sind häufig Wortfindungsstörungen bzw. Sprachstörungen (Aphasien) bei Patienten mit einem schweren Schädel-Hirn-Trauma zu beobachten [55, S. 9].

Bei Menschen mit einer Intelligenzminderung hängen die Auswirkungen auf die Interaktionsfähigkeiten vom Grad der Einschränkung ab [72, S. 222]. So haben Menschen mit einer leichten Intelligenzminderung häufig Probleme sich sprachlich auszudrücken, können die Folgen einer Handlung nur begrenzt abschätzen und haben eine geringere Frustrationsgrenze. In Fällen einer schwereren Intelligenzminderung sind körperliche und neurologische Behinderungen meist so stark ausgeprägt, dass eine Interaktion mit einem System unmöglich wird.

3.1.3.4 Verteilungen von Einschränkungen in der Bevölkerung

Es existieren verschiedene Arten von Demenz, wobei die Alzheimer-Demenz mit 72 %, die Vaskuläre Demenz mit 15 % und die Parkinson-Demenz mit 6 % die drei häufigsten Formen sind. In Tabelle 8 sind die Häufigkeiten (Prävalenzraten) für Demenzerkrankungen in Deutschland und Europa aufgeführt. Wie aus den Daten ersichtlich ist, sind insbesondere ältere Menschen von Demenz betroffen.

In Deutschland ereignen sich jährlich rund 262.000 Schlaganfälle, wobei ein Großteil der Personen über 70 ist und rund 25 % der Patienten versterben [99]. Etwa 60 % der Patienten, die einen Schlaganfall überleben, haben bleibende neurologische Schäden [192, S. 27], allerdings können 60 bis 75 % der Patienten nach einer längeren Rehabilitationsphase selbstständig leben [240]. Von einem Schädel-Hirn-Trauma (SHT) sind rund 200.000 Personen pro Jahr betroffen [153, S. 577]. Rund ein Drittel erleidet ein schweres SHT. Während leichte SHT meist vollständig heilen oder leichte Einschränkungen auftreten, bleiben bei schweren SHT etwa 10 % der Patienten schwerstbehindert [195, S. 1074].

In der Tabelle 9 sind Häufigkeiten für kognitive Beeinträchtigungen aufgeführt, die nach einem Schädel-Hirn-Trauma (SHT) oder Schlaganfall auftreten können. Die in der Tabelle mit ^a markierten Werte sind aus der Studie von Tuomilehto et al. [235] entnommen, die Schlagan-

Altersgruppe	Maercker [159, S. 60]	Ritchie & Kildea [191, S. 933]	Hofman et al. [101, S. 742]
30 - 59 Jahre	-	-	0,1 %
60 - 64 Jahre	0,5 %	-	1,0 %
65 - 69 Jahre	0,9 %	1,53 %	1,4 %
70 - 74 Jahre	2,6 %	3,54 %	4,1 %
75 - 79 Jahre	6,1 %	6,8 %	5,7 %
80 - 84 Jahre	14,8 %	13,57 %	13 %
85 - 89 Jahre	-	22,26 %	21,6 %
90 - 94 Jahre	-	31,28 %	32,2 %
über 95 Jahre	-	44,28 %	34,7 %

Tabelle 8: Prävalenzraten für Demenzerkrankungen

Probleme nach ...	Schädel-Hirn- Trauma	Schlaganfall
Gedächtnisprobleme	54 % ^b	-*
Schnellere Ermüdung	30 % ^b	-*
Irritierbarkeit	39 % ^b	-*
Langsamkeit	33 % ^b	70 % ^c
Aufmerksamkeitsprobleme	33 % ^b	-
Kein Multitasking möglich	21 % ^b	-*
Lähmung (leicht)	-*	47 % ^a
Lähmung (schwer)	-*	11,5 % ^a
Sprachstörung	-*	29,5 % ^a

Tabelle 9: Häufigkeiten für Einschränkungen.

* keine Werte verfügbar

fallpatienten rund 14 Jahre nach dem Ereignis interviewt haben. Die mit ^b markierten Werte sind subjektiv erlebte Häufigkeiten, wie sie zwei Jahre nach einem SHT von Patienten berichtet wurden [261]. Die Werte die mit ^c markiert wurden, sind bei Lelow und Paetow [153, S. 577] zu finden.

Die Prävalenz von Intelligenzminderung liegt in Industriestaaten zwischen 3 und 4 % [72, S. 217]. Die Intelligenzminderung wird hinsichtlich ihrer Schwere in vier Kategorien (*leichte, mittelschwere, schwere* und *schwerste Intelligenzminderung*) klassifiziert [255], wobei die leichte Minderung (IQ 50 - 69) mit einem Anteil von rund 80 % die am häufigsten zu findende Form ist [89, S. 15].

3.1.3.5 Zusammenfassung

Nach Persad et al. [178] sind die Kapazität des Kurzzeit- und Arbeitsgedächtnisses und die Schnelligkeit und Genauigkeit der Verarbeitung wichtig, um optimal mit einem System interagieren zu können. Wie in diesem Abschnitt gezeigt wurde, existieren verschiedene Einschränkungen, die die Kognition limitieren und sich somit auf die Interaktion auswirken.

3.1.4 Motorik

In diesem Abschnitt werden Einschränkungen beschrieben, die die Ausführungen von Interaktionen limitieren oder verhindern. Insbesondere wird auf krankhafte Bewegungsstörungen, wie den Tremor, das Parkinson-Syndrom oder die rheumatischen Erkrankungen, und auf erworbene Körperbehinderungen eingegangen. Die Auswahl der Einschränkungen basiert auf der Verfügbarkeit von Daten zur Prävalenz in der Bevölkerung.

3.1.4.1 Überblick

Der Begriff *Motorik* beschreibt die vom Gehirn koordinierten Bewegungsabläufe und kann in Grob- und Feinmotorik untergliedert werden [184]. Der Begriff der Grobmotorik umfasst Bewegungen, die den gesamten Körper umfassen und mit Kraft ausgeübt werden, während die Feinmotorik genaue Bewegungen umfasst, die in der Regel nur einen geringen Kraftaufwand benötigen. Viele körperliche Attribute (z. B. Körpergröße, Gewicht und Armreichweite) sind durch genetische, ethnische und soziale Faktoren beeinflusst, allerdings können diese durch Krankheiten und Unfälle eingeschränkt werden. Nachfolgend werden verschiedene krankhafte und erworbene Einschränkungen erläutert.

3.1.4.2 Risikofaktoren und Ursachen für Einschränkungen

Der Tremor ist eine unwillkürliche, rhythmische und oszillierende Bewegung die meistens die Hand betrifft [51, S. 255]. Im Allgemeinen werden vier Arten unterschieden: der Ruhetremor, der Haltetremor, der essenzielle Tremor und der Intentionstremor (zerebellärer Tremor) [175, S. 13]. Die Anlagen für einen Tremor werden meist vererbt, allerdings ist unklar, was die Krankheit auslöst [53, S. 1401]. Der essenzielle Tremor hat meist eine Frequenz von 6 bis 7 Hertz (Hz) und wird durch psychische Belastungen und Aufregung verstärkt [53, S. 1410]. Die Verbreitung des essenziellen Tremors nimmt ab dem 65. Lebensjahr deutlich zu, ein Unterschied zwischen den Geschlechtern ist nicht festzustellen [185, S. 191]. Ein Ruhetremor hat eine Frequenz von 4 bis 6 Hz [53, S. 1410]. Er tritt nur in Ruhe auf und wird durch Bewegung gehemmt. Der Ruhetremor ist häufig ein Symptom bei der Parkinson-Krankheit [185, S. 208]. Der Haltetremor hat meist eine zerebelläre (neuronale) Schädigung als Ursache und kann insbesondere bei Alkoholikern beobachtet werden [51, S. 262]. Beim Intentionstremor beträgt die Frequenz 3 bis 5 Hz [53, S. 1410]. Die Ursache für diesen Tremor ist mit hoher Wahrscheinlichkeit eine Verletzung oder Funktionsstörung im Kleinhirn [51, S. 272].

Das Parkinson-Syndrom schränkt die Ausführung von unwillkürlichen und willkürlichen Bewegungsabläufe ein [57, S. 433]. In einigen Veröffentlichungen wird berichtet, dass das Parkinson-Syndrom häufiger Männer als Frauen (Männer : Frauen = 1,46 : 1) betrifft [50, S. 36], in anderen Publikationen findet sich dieser Unterschied nicht [33, S. 35]. Die Ursachen für das Parkinson-Syndrom sind weitestgehend unbekannt [50, S. 37]

Unter dem Begriff *rheumatische Erkrankungen* sind Erkrankungen wie Arthrose, rheumatoide Arthritis und Gicht zusammengefasst [15, S. 172]. Bei der Arthrose handelt es sich um eine Gelenkerkrankung, die durch eine Zerstörung des Gelenks und durch Schädigung der umliegenden Strukturen charakterisiert ist [183, S. 7]. Bei der primären Arthrose ist die Ursache nicht eindeutig, als Risikofaktoren gelten Alter, weibliches Geschlecht, genetische Faktoren, hohe Belastungen (schwere Arbeit und insbesondere Übergewicht) [183, S. 15 ff] und auffallend, ein niedriger Bildungsstatus bei den Frauen [148, S. 91]. Die rheumatoide Arthritis ist eine Erkrankung, die meist ältere Personen betrifft und mit 55 bis 65 Jahren beginnt, wobei Männer meist später erkranken [260, S. 8 ff]. Die Anlagen für eine rheumatoide Arthritis werden vererbt, allerdings existieren noch weitere, meist unklare Ursachen, die die Erkrankung auslösen. Als ein Risikofaktor gilt das Rauchen; je mehr und je öfter geraucht wird, desto häufiger tritt die Krankheit auf. Die Gichterkrankung tritt meist in Zusammenhang mit krankhaftem Übergewicht (Adipositas), Diabetes mellitus und Bluthochdruck auf [230]. Die Ursachen für eine Gichterkrankung werden vererbt, allerdings spielen auch die Ernährung und der Alkoholkon-

sum (insbesondere Bier) eine Rolle. Von einer Gichterkrankung sind meist ältere Menschen betroffen und überdies Männer häufiger als Frauen. In einer aktuellen Studie weisen Tausche et al. [229] darauf hin, dass das Risiko einer Gichterkrankung steigt, wenn verschiedene Krankheiten durch unterschiedliche Medikamente behandelt werden, wie es im Alter häufig passiert.

Die Ursache für eine Körperbehinderung kann vielfältigen Ursprungs sein. Im Allgemeinen wird zwischen einer angeborenen Behinderung, einer Behinderung nach einem Unfall und einer Behinderung durch eine Krankheit unterschieden. Nach der aktuellen Statistik des Landes Brandenburg [2], sind rund 88 % der Menschen mit einer schweren Behinderung durch die Folgen einer Erkrankung behindert, rund 2 % sind aufgrund eines Unfalls dauerhaft eingeschränkt, 4,5 % beträgt der Anteil derer, die von Geburt an eine Behinderung haben und bei 5,5 % der Menschen liegen mehrere oder unklare Ursachen vor.

Unter *vorübergehenden Einschränkungen* werden in dieser Arbeit *Unfallverletzungen* verstanden, die keine dauerhaften Einschränkungen verursachen. Bei den temporären Einschränkungen sind Männer, insbesondere jüngere Männer, stärker als Frauen betroffen [149, S. 58]. Mit dem Alter gleicht sich dieser Unterschied aus, um bei den über 65-Jährigen zu verschwinden.

3.1.4.3 Auswirkungen

Eine Einschränkung in der Motorik führt meist zu Problemen in der Interaktion mit einem System. Dabei sind die resultierenden Probleme sehr unterschiedlich und von der Art der Einschränkung abhängig.

Beim essenziellen Tremor tritt das Problem auf, dass zielgerichtete Eingaben nur schwer oder gar nicht zu tätigen sind. So treten bei einem Touchscreen, je nach Schwere der Ausprägung, große Probleme in der Ausführung der intendierten Interaktion auf. Da der essenzielle Tremor durch Stress verstärkt wird, können durch Interaktionsfehler und Zeitdruck die Symptome verschlimmert werden, was die Interaktion zusätzlich einschränkt. Beim Ruhe-Tremor gibt es in der aktiven Bewegung keine oder nur geringe Probleme. Allerdings können Probleme durch unwillkürliches Auslösen von Interaktionen in Ruhe auftreten. Der zerebelläre Tremor ist problematisch in der Interaktion, da große Probleme bei der Umsetzung von intendierten Interaktionen entstehen, die die Ausführung der Interaktion stark behindern.

Das Parkinson-Syndrom wirkt sich stark auf die Bewegung aus. Bei allen erkrankten Personen ist eine Verlangsamung der Bewegungen zu beobachten. Immerhin 75 % der Erkrankten haben einen Ruhetremor (bei den meisten Patienten zusätzlich ein Haltetremor [51, S. 261]), 64 % entwickeln Haltungsschäden und 49 % der Betroffenen haben Sprechstörungen [50, S. 47].

Die rheumatischen Erkrankungen wirken sich oftmals auf die Beweglichkeit einzelner Gliedmaße aus. Eine Arthrose führt aufgrund von Schmerzen zu Einschränkungen in der Bewegung [183, S. 7]. Bei einer rheumatoiden Arthritis hängen die Einschränkungen oft davon ab, wann mit der Behandlung begonnen wurde. Bei einem frühen Behandlungsbeginn sind rund 50 % der Personen beschwerdefrei [260, S. 8 ff]. Treten Beschwerden auf, so sind die betroffenen Personen durch Schmerzen und Bewegungseinschränkungen in ihrer Beweglichkeit stark limitiert. Bei der Gichterkrankung sind es insbesondere, wie bei den anderen rheumatischen Erkrankungen, die Schmerzen, die die Beweglichkeit einschränken und so die Interaktion mit dem System beeinträchtigen.

Eine Körperbehinderung wirkt sich in Abhängigkeit vom Grad und der Art der Behinderung auf die Interaktion aus. Fehlende obere Extremitäten können z. B. die Ausführung von Mehrfingergesten auf einem Touchscreen behindern, da das Gerät nicht gehalten werden kann und eventuell nicht stabil liegt. Fehlende untere Extremitäten können z. B. die Erreichbarkeit von zu hoch platzierten Eingabeobjekten (Button und Knöpfe an Automaten) einschränken.

Die Auswirkungen von temporären Einschränkungen sind vielfältig. Bei den häufigsten Verletzungsarten (Verrenkungen, Verstauchungen, Zerrungen und Bänderrisse) wird die Beweglichkeit der betroffenen Körperteile für mehrere Wochen stark eingeschränkt. Bei Wunden, oberflächlichen Verletzungen und Quetschungen ist die Auswirkung auf die Interaktion schwierig zu beurteilen, da die Folgen stark von der Art, Schwere und Lokalisation der Verletzung abhängen.

3.1.4.4 *Verteilungen von Einschränkungen in der Bevölkerung*

Zur Bestimmung der Normalwerte der Körpermaße von deutschen Bundesbürgern im Alter zwischen 18 und 65 Jahren wird auf die DIN 33402 und dort auf den Teil 2 [48] verwiesen. Für jüngere und ältere Menschen existieren keine genormten anthropometrische Daten, allerdings kann für Kinder und Jugendliche auf die KiGGS-Studie des Robert Koch Institutes [167] verwiesen werden, die zumindest für einige Attribute (z. B. Körpergröße und Gewicht), die entsprechenden Angaben enthält. Für Menschen über 65 Jahren existieren keine vergleichbaren Normen, so dass auf vereinzelte Studien [130] und Dissertationen [22, 248] zurückgegriffen werden muss.

Beim Tremor gibt es Unterschiede in der Verteilung zwischen den verschiedenen Arten. Der essenzielle Tremor hat eine Prävalenz von 1,3 bis 5,1 % [185, S. 191], wobei die Spannbreite aus den verschiedenen Altersgruppen resultiert. Beim essenziellen Tremor sind verschiedene Körperbereiche unterschiedlich oft betroffen. Bei 94 % der Patienten sind die Hände betroffen, 33 % der Personen haben einen Tremor, der den Kopf betrifft, 16 % der Patienten haben Probleme mit der Stimme, 8 % mit dem Kinn, 3 % mit dem Gesicht und 2,5 % der Patienten haben

Altersgruppe	Frauen	Männer
18 - 29 Jahre	2,0 %	1,4 %
30 - 44 Jahre	7,4 %	6,1 %
45 - 64 Jahre	28,4 %	21,6 %
ab 65 Jahren	46,9 %	30,5 %

Tabelle 10: 12-Monats-Prävalenz von Arthrose bei Männern und Frauen in verschiedenen Altersgruppen [193, S. 83]

einen Tremor, der den Rumpf betrifft [185, S. 195]. Ein Ruhetremor tritt meist nur mit einer Parkinsonerkrankung auf. Der zerebelläre Tremor kommt deutlich seltener vor.

Das Parkinson-Syndrom betrifft 0,1 % bis 0,2 % der Gesamtbevölkerung, allerdings erhöht sich der Anteil mit zunehmenden Alter, so sind es in der Bevölkerung über 65 Jahren schon 1 bis 2 % betroffen [50, S. 35]. Vereinzelt wurden schon deutlich höhere Fallzahl (über 10 % der Personen zwischen 65 und 74 Jahren und über 40 % bei Personen über 85 Jahren) berichtet [33, S. 35].

Die Prävalenz für Arthrose liegt bei 28 % für Frauen und bei 20 % für Männer [193, S. 82]. Demgegenüber haben nur etwa 0,5 bis 0,8 % der Gesamtbevölkerung eine rheumatoide Arthritis [260, S. 8] und 1,4 % eine Gichterkrankung [230]. Die Arthrose betrifft hauptsächlich ältere Menschen. Bei den unter 30-jährigen Personen sind nur etwa 1,6 % betroffen. Bei den 30 bis unter 50-Jährigen steigt die Prävalenz auf durchschnittlich 14,9 % (16,6 % Frauen und 13,3 % Männer). Zwischen 50 und 60 Jahren gibt es einen weiteren starken Anstieg; hier sind rund 33 % bei den Frauen und 25 % bei den Männern betroffen. Ab 60 Jahren hat etwa jede zweite Frau und jeder dritte Mann eine Arthrose [183, S. 15]. Allerdings wirken sich nicht alle Arthrosen gleich auf das Interaktionsverhalten aus. Eine Arthrose betrifft meist die Knie- (48,4 %) oder Hüftgelenke (39,0 %) und seltener andere Gelenke [183, S. 18]. Nach Engelhardt [61], haben allerdings rund 30 % der betroffenen Menschen eine Arthrose im Daumen oder Fingergelenk; andere Studien [197] sprechen sogar von einem Anteil von bis zu 76 % in den Fingerendgelenken. In Tabelle 10 ist die 12-Monats-Prävalenz für Männer und Frauen in unterschiedlichen Altersklassen zusammengefasst. Die 12-Monats-Prävalenz beschreibt den Anteil der Personen, die innerhalb von 12 Monaten mindestens einmal von der entsprechenden Krankheit betroffen waren.

In Deutschland leben rund 7,3 Millionen Menschen (8,9 %) die als Schwerbehinderte anerkannt sind [107, S. 8]. Von diesen 7,3 Millionen Menschen sind 29,2 % über 75 Jahre alt; 45,8 % gehören zu Altersgruppe zwischen 55 und 75 Jahren. Bei den jüngeren Menschen beträgt

Schwerste Behinderung	Frauen	Männer
Verlust oder Teilverlust mindestens eines Armes	0,2 %	0,4 %
Funktionseinschränkung von mindestens einem Arm	3,1 %	3,2 %

Tabelle 11: Häufigkeiten der schwersten Behinderung nach [107, S. 22 ff]

Altersgruppe	Frauen	Männer
18 - 29 Jahre	10,5 %	19,4 %
30 - 44 Jahre	6,6 %	13,0 %
45 - 64 Jahre	6,1 %	8,6 %
ab 65 Jahren	5,0 %	5,4 %

Tabelle 12: Unfallverletzungen pro Altersgruppe nach [149, S. 59]

der Anteil der Menschen mit Behinderung 4,0 %. In Tabelle 11 sind Gründe für eine Schwerbehinderung mit Einfluss auf die Motorik und deren Häufigkeiten aufgeführt.

Rund 9 % der Erwachsenen erleiden pro Jahr einen nicht-tödlichen Unfall, der durch einen Arzt behandelt werden muss [149, S. 58]. In der Tabelle 12 sind die Häufigkeiten hierfür pro Altersgruppe spezifiziert. Wie in der Tabelle dargestellt, sind besonders die jüngeren Männer (18 bis 29 Jahre) betroffen. Hier haben rund 19 % einen Unfall, der von einem Arzt behandelt werden muss. Ab einem Alter von 65 Jahren, ist das Verhältnis zwischen den Geschlechtern ausgeglichen. Bei rund 1,6 % der verunfallten Personen, war die Verletzung so schwerwiegend, dass eine beantragte Behinderung zuerkannt wurde [237, S. 19].

3.1.4.5 Zusammenfassung

In diesem Abschnitt wurde dargestellt, welche motorischen Einschränkungen einen Einfluss auf das Interaktionsverhalten haben, welche Ursachen diese haben und wie die Verteilungen in der Bevölkerung sind. Eine Einschränkung in der Motorik resultiert häufig aus einer Erkrankung, einer Behinderung oder aus einer Unfallverletzung, allerdings wird der Anteil der Einschränkungen, die aus einer Erkrankung resultieren weiter zunehmen, da der demografische Wandel zu einem höheren Anteil älterer Menschen führt.

3.1.5 Weitere Attribute

In diesem Abschnitt werden Krankheiten und andere Faktoren betrachtet, die in keine der vorherigen Kategorien passen, die allerdings einen direkten oder indirekten Einfluss auf die Wahrnehmung, Kognition oder Motorik haben. Da die Anzahl der möglichen Faktoren sehr groß ist, werden nur wenige, besonders relevante Ursachen betrachtet, hierzu zählen Bluthochdruck, Diabetes mellitus, starkes Übergewicht (Adipositas), Depression und Multiple Sklerose. Alle in diesem Abschnitt beschriebenen Attribute zählen, mit Ausnahme der Multiplen Sklerose, zu den häufigen Erkrankungen in Deutschland.

3.1.5.1 Risikofaktoren und Ursachen

Der Bluthochdruck, auch Hypertonie genannt, ist ein wichtiger Risikofaktor, der zu einer Herz-Kreislauf-Erkrankung führen kann [193, S. 25]. Als Risikofaktoren für die Entwicklung einer Hypertonie gelten das Alter, das Geschlecht, eine ungesunde Ernährung (viel Kochsalz und hoher Alkoholkonsum), viel Stress und zu wenig Bewegung.

Beim Diabetes mellitus handelt es sich um eine Stoffwechselerkrankung die in zwei Formen unterschieden wird [193, S. 65]. Der Typ-1-Diabetes tritt meist bereits bei Kindern und Jugendlichen auf. Hierbei handelt es sich um eine Autoimmunerkrankung, die die Zellen zerstört, die das Insulin produzieren. Beim Typ-2-Diabetes, auch Altersdiabetes genannt, liegt kein Insulinmangel vor, sondern eine verminderte Insulinempfindlichkeit [91, S. 4]. Zu den Risikofaktoren für den Typ-1-Diabetes gelten genetische Faktoren und Umweltfaktoren (Ernährung, Virusinfektion, Stress, sozioökologische Faktoren), allerdings sind die ursächlichen Zusammenhänge noch nicht vollkommen geklärt [91, S. 5]. Der Typ-2-Diabetes ist ebenfalls durch erbliche Faktoren beeinflusst, allerdings gibt es verschiedene verhaltensbedingte Faktoren, die das Risiko erhöhen. Zu den verhaltensbedingten Faktoren zählten Übergewicht, körperliche Inaktivität, Rauchen und das metabolische Syndrom (eine Kombination verschiedener Risikofaktoren) [91, S. 5 ff].

Eine Adipositas entsteht, wenn über eine längere Zeit mehr Nahrungsenergie zugeführt wird, als der Körper am Tag verbraucht [24, S. 23]. Dabei ist ein starkes Übergewicht (Adipositas), anders als oftmals vermutet, zum Großteil erblich bedingt, da die Vererbung den Energieverbrauch beeinflusst [253, S. 30]. Durch Fehlernährung und Bewegungsmangel wird das Energieungleichgewicht verstärkt, was das Risiko einer Adipositas erhöht [24, S. 27 ff]. Allerdings sind die genetischen Faktoren, das Essverhalten und die Bewegungsarmut nicht immer die ausschlaggebenden Risikofaktoren. So konnte in verschiedenen Studien gezeigt werden, dass die Schwangerschaft in der Altersgruppe der 25 bis 34-jährigen Frauen ein Hauptrisikofaktor für die Entstehung einer Adipositas ist [138, S. 72].

Bei der Depression handelt es sich um eine psychische Störung, die unterschiedliche Schweregrade aufweisen kann [193, S. 69]. Für eine Depression kommen unterschiedliche Ursachen in Betracht; neben körperlichen Erkrankungen, medikamentösen Auslösern, sozialen Gründen können vor allem auch erbliche Faktoren eine Rolle spielen [64].

Die letzte Erkrankung auf die in diesem Abschnitt eingegangen wird ist die Multiple Sklerose (MS). Bei einer MS-Erkrankung wird das zentrale Nervensystem durch chronische Entzündungen geschädigt. Die Ursachen sind unbekannt, es wird vermutet, dass Umwelteinflüsse, Erbanlagen und Fehlreaktionen des Immunsystems zur Erkrankung führen [140, S. 26]. Frauen sind häufiger von MS betroffen als Männer [256, S. 2].

3.1.5.2 Auswirkungen

Die in diesem Abschnitt beschriebenen Erkrankungen haben meist keine direkte Auswirkung auf die Interaktion, beeinflussen allerdings die in den vorherigen Abschnitten genannten Erkrankungen und haben hierdurch zumindest einen indirekten Einfluss auf die Interaktion.

Durch Bluthochdruck steigt die Gefahr einen Schlaganfall zu erleiden um das Zwei- bis Vierfache [190]. Eine Hypertonie beeinflusst allerdings auch die Sinnesorgane. So kann durch einen hohen Blutdruck das Auge geschädigt und die Sehfähigkeit reduziert werden bzw. ganz verloren gehen [156].

Bei Vorliegen eines Diabetes mellitus besteht ein höheres Risiko für Herz-Kreislauf-Erkrankungen (Herzinfarkt), Schlaganfall, Nierenerkrankungen (Nephropathie), Augenerkrankungen (Retinopathie), Nervenerkrankungen (Neuropathie) und Amputationen (diabetische Fußsyndrom). Weiterhin ist das Herzinfarktrisiko bei Männern 3,9-mal und bei Frauen 5,9-mal höher. Für einen Schlaganfall steigt das Risiko um das zwei bis dreifache. Werden Augenerkrankungen betrachtet, so zeigt sich für Diabetiker ein 5-mal höheres Risiko zu erblinden [91, S. 10 ff].

Eine Adipositas ist meist ursächlich für verschiedene andere Krankheiten bzw. erhöht das Risiko für diese Krankheiten. Bei übergewichtigen Personen können häufig Herz-Kreislauf-Erkrankungen, Diabetes mellitus, Fettlebererkrankungen, Fettstoffwechselstörungen, Bluthochdruck, Erkrankungen des Verdauungstraktes, Atembeschwerden und Beeinträchtigungen des Bewegungsapparates vorliegen [254, S. 173 ff].

Bei einer Depression treten verschiedene Krankheitszeichen auf. Neben dem auffälligsten Merkmal, der Stimmungsänderung, treten verschiedene andere Beschwerden auf. So kann durch Denk- und Konzentrationsstörungen, verlangsamter Reaktionsfähigkeit, Bewegungsstörungen und starke Ermüdungserscheinungen die Interaktion mit einem System eingeschränkt werden [64].

Altersgruppe	Frauen	Männer
18 - 29 Jahre	1,0 %	0,9 %
30 - 44 Jahre	1,5 %	1,6 %
45 - 64 Jahre	6,5 %	9,4 %
ab 65 Jahren	17,4 %	18,6 %

Tabelle 13: Verteilung für die 12-Monats-Prävalenz von Diabetes mellitus [193, S. 66]

Komplikation	Frauen	Männer
Herzkomplikationen	26,0 %	20,6 %
Diabetische Augenerkrankung	15,5 %	13,9 %
Diabetisches Nervenleiden	11,8 %	9,2 %
Amputation	1,4 %	2,6 %

Tabelle 14: Häufigkeiten von Komplikationen bei Diabetes mellitus; ausgewählte Komplikationen nach [97]

Die Auswirkungen der Multiplen Sklerose sind, aufgrund der Erkrankung des zentralen Nervensystems, vielfältig und hängen von Krankheitsverlauf ab. Bei der Erstdiagnose weisen viele Patienten (> 30 %) Sensibilitätsstörungen, Einschränkungen in der visuellen Wahrnehmung und Lähmungserscheinungen auf. Die Einschränkungen in der visuellen Wahrnehmung äußern sich dadurch, dass die betroffenen Personen nur noch verschwommen oder milchig sehen können. Im Laufe der Erkrankung entwickeln 80 bis 90 % der Patienten eine Sehstörung während eines Schubes, die Sehkraft kann hierbei zeitweise stark abnehmen, eine vollständige dauerhafte Erblindung kommt allerdings nur sehr selten vor [256, S. 4].

3.1.5.3 Verteilungen von Einschränkungen in der Bevölkerung

Nach einer Studie von Haerting et al. [84] haben rund 74,3 % der Männer und 68,4 % der Frauen ab 45 Jahren eine Hypertonie. Werden die Zahlen nach Altersgruppen aufgeschlüsselt, so zeigt sich, dass der Anteil der betroffenen Menschen mit dem Alter steigt. Sind es bei den 45 bis 54-jährigen Teilnehmern der Studie noch knapp über 50 % (Frauen: 55,4 %; Männer: 58,7 %), so vergrößert sich der Anteil bei den 75 bis 83-Jährigen auf über 80 % (Frauen: 86,5 %; Männer: 83,6 %).

Beim Diabetes mellitus muss zwischen den beiden Diabetes-Typen unterschieden werden. Vom Typ-1-Diabetes sind etwa 0,7 % der Bevölkerung betroffen. Der Typ-2-Diabetes oder Altersdiabetes kommt

Altersgruppe	Übergewichtig		Adipositas	
	Frauen	Männer	Frauen	Männer
18 - 29 Jahre	30,0 %	35,3 %	9,6 %	8,6 %
30 - 39 Jahre	38,0 %	62,4 %	17,9 %	22,0 %
40 - 49 Jahre	46,4 %	69,9 %	18,6 %	22,9 %
50 - 59 Jahre	60,9 %	78,3 %	27,3 %	27,5 %
60 - 69 Jahre	70,7 %	83,9 %	34,8 %	33,1 %
70 - 79 Jahre	80,3 %	82,5 %	41,6 %	31,3 %

Tabelle 15: Übergewicht und Adipositas in den Altersgruppen nach [161]

häufiger vor, hiervon sind etwa 11,1 % in der Bevölkerung betroffen. Das Verhältnis von Typ-1 und Typ-2-Diabetikern beträgt etwa 1 zu 9 [91, S. 4]. In der Tabelle 13 ist die Häufigkeitsverteilung für das Vorliegen eines Diabetes mellitus in den verschiedenen Altersgruppen beschrieben. Wie der Tabelle zu entnehmen ist, steigt die Prävalenz mit dem Alter an. Im Zusammenhang mit Diabetes kommt es häufig zu Komplikationen, in Tabelle 14 sind deren Häufigkeiten aufgeführt. Wie zu erkennen ist, sind Herzkomplikationen sehr häufig, ebenso Augenerkrankungen und Nervenleiden.

Nach einer aktuellen Studie [161] sind in Deutschland 67,1 % der Männer und 53,0 % der Frauen übergewichtig. Wird der Anteil der Personen mit starkem Übergewicht betrachtet, so zeigt sich, dass etwa jeder vierte Mann (23,3 %) und jede vierte Frau (23,9 %) stark übergewichtig (adipös) ist. In Tabelle 15 sind die Verteilungen für Übergewicht und Adipositas in den verschiedenen Altersgruppen dargestellt. Wie aus der Tabelle ersichtlich, nimmt der Anteil von übergewichtigen und adipösen Menschen mit dem Alter zu. Ist in der jüngeren Altersgruppe nur jeder Dritte übergewichtig, so sind in der ältesten Altersgruppe bereits 4 von 5 Menschen übergewichtig und jeder Dritte stark übergewichtig. Weiterhin hat die Zugehörigkeit zu einer sozialen Schicht einen starken Einfluss auf das Gewicht, so sind die Angehörigen einer niedrigen Schicht eher von Übergewicht und Adipositas betroffen, als die Personen mit einem hohen sozialen Status.

Bei der Depression ist der Anteil derer, die akut an einer Erkrankung leiden mit etwa 10 % relativ hoch. Werden nur die schweren Depressionen betrachtet, so liegt der Anteil noch immer bei 2 bis 7 % [64]. In verschiedenen Altersgruppen zeigt sich ein signifikanter Geschlechter- und Bildungsunterschied bei der Prävalenz, so sind in der Gruppe der Personen über 45 Jahren mehr Frauen als Männer von der Krankheit betroffen und bei den Personen über 65 Jahren sind

Frauen aus den unteren sozialen Schichten doppelt so häufig belastet. Bei den Männern ist der Unterschied in der Altersgruppe von 45 bis 64 Jahren zwischen unterer und oberer Bildungsgruppe besonders groß, hier haben die Männer aus der unteren Bildungsgruppe ein 3-mal höheres Depressionsrisiko [193, S. 69].

Die Multiple Sklerose kommt in erster Linie in westlichen Industrienationen vor. Die Prävalenz liegt hier bei 100 bis 120 Personen pro 100.000 Einwohnern [256, S. 2], was relativ niedrig erscheinen mag, allerdings zählt die MS zu den häufigsten Gründen für eine Invalidität im jungen und mittleren Lebensalter [160, S. 351]

3.1.5.4 Zusammenfassung

In diesem Teil der Arbeit wurden verschiedene Erkrankungen vorgestellt, die keinem der vorherigen Abschnitte zugeordnet werden konnte, da sie entweder keinen direkten Einfluss auf das Interaktionsverhalten haben oder in ihren Auswirkungen mehrere Bereiche betreffen.

3.2 MODELLIERUNG VON NUTZERATTRIBUTEN

In diesem Teil der Arbeit wird beschrieben, wie die in Abschnitt 3.1 dargestellten Nutzerattribute im Modell abgebildet werden können. Zu Beginn dieses Abschnitts wird in einer Einführung beschrieben, wieso ein Bayes'sches Netz zur Abbildung der Nutzerdaten verwendet wird, wie Bayes'sche Netze funktionieren und wofür sie im Allgemeinen genutzt werden. Im Anschluss an diese Einführung wird die Struktur des Bayes'schen Netzes dargestellt, dass die Nutzerattribute repräsentiert und das zum Generieren der demografischen Nutzermodelle verwendet wird.

3.2.1 Einführung in Bayes'sche Netze

Eine kurze Einführung in das Thema kann weder vollständig noch sonderlich tiefgreifend sein, deshalb wird für die vertiefende Darstellung auf das Buch von Jensen und Nielsen [118] verwiesen, die eine gut verständliche und theoretisch fundierte Einführung in das Thema geben. Die nachfolgende Ausführung fassen die wichtigsten Fakten über Bayes'sche Netze aus dem genannten Buch zusammen.

3.2.1.1 Motivation

Das Wissen über Nutzer und deren Einschränkungen kann in vielfältiger Weise gespeichert werden. In Abschnitt 1.4.2 wurde beschrieben, dass für die Bereitstellung von Daten bisher Datenbanken (HADRIAN) und Ontologien (BS) verwendet werden. Beide Lösungen bieten Vor- und Nachteile gegenüber einem Bayes'schem Netz. Ein relationales

Datenbanksystem eignet sich gut, um strukturierte Informationen abzulegen. Die Zugriffsgeschwindigkeit ist meist hoch und das Anlegen einer Datenbank sehr einfach. Probleme entstehen, wenn oft gleiche Informationen gespeichert werden sollen. Durch *Normalisierung* kann Redundanz vermieden werden, allerdings geschieht dies auf Kosten der Geschwindigkeit und Übersichtlichkeit. Ein weiterer Nachteil einer datenbankbasierten Lösung ist, dass die Zusammenhänge oder Abhängigkeiten zwischen den einzelnen Informationen nicht gespeichert werden.

Eine *Ontologie* beschreibt, auf einem abstrakten Level, verschiedene Konzepte (Klassen) und deren Zusammenhänge. Die *Instanz* einer Ontologie ist eine konkrete Ausprägung der definierten Konzepte und Zusammenhänge. Mit Hilfe einer Ontologie können die in Abschnitt 3.1 beschriebenen Nutzerattribute und deren Zusammenhänge modelliert werden. Verschiedene Algorithmen erlauben anhand der Beschreibung der Ontologie fehlende Informationen (logische Schlüsse) zu berechnen und so implizites Wissen explizit zu machen. Ein Nachteil der Ontologie ist, dass Ontologien kein unsicheres Wissen verarbeiten können. Wenn die Information über die Ausprägung einer Variable nicht sicher ist, können in einer auf einer Ontologie basierten Lösung keine Schlüsse gezogen werden.

Ein Bayes'sches Netz bildet Zusammenhänge zwischen Variablen effizient ab und kann auch Unsicherheit verarbeiten, weshalb dieser Ansatz in der vorliegenden Arbeit gewählt wurde.

3.2.1.2 Theorie

Bei einem Bayes'schen Netz (BN) handelt es sich um einen gerichteten azyklischen Graphen, der aus *Knoten* und *Kanten* besteht. Die Knoten in einem BN repräsentieren Variablen und die Kanten markieren Zusammenhänge zwischen je zwei Variablen. Ein Knoten hat eine Menge von Ausprägungen, man spricht in diesem Fall vom *Sample space* (SP). Der SP enthält alle möglichen, sich allerdings gegenseitig ausschließende, Zustände einer Variable und beschreibt diese somit vollständig. Eine Teilmenge des SP wird als *Event* (A) definiert und kann mathematisch als $A \subseteq SP$ beschrieben werden. Ein Event kann ein oder mehrere Ausprägungen der Variable umfassen und jedem Event ist eine *Wahrscheinlichkeit* $P(A)$ zugeordnet. Zur Verdeutlichung soll das nachfolgende Beispiel dienen.

Ein Würfel kann sechs verschiedene Werte anzeigen, der *Sample space* ist definiert als $SP_W = \{1, 2, 3, 4, 5, 6\}$. Durch das Würfeln wird ein zufälliges *Ergebnis*, z. B. $a_1 = 3$, erzielt. Wenn angenommen wird, dass es sich um einen fairen Würfel handelt, so ist die Wahrscheinlichkeit, dass eine der Zahlen aus dem SP gewürfelt wird, für alle Elemente gleich: $P(a_1) = P(a_2) = \dots = P(a_6)$. Da die Summe der Wahrscheinlichkeiten über alle Ergebnisse 1 betragen muss, ist die Wahrscheinlichkeit des Auftretens von a_1 durch $P(a_1) = \frac{1}{6}$ definiert.

Handelt es sich um einen unfairen Würfel, der z. B. durch das Anlegen eines magnetischen Feldes beeinflusst wird, so kann dieser Umstand ebenfalls durch eine Wahrscheinlichkeitsverteilung beschrieben werden; nur sind die Wahrscheinlichkeiten für die möglichen Ergebnisse nicht mehr gleich verteilt. Ein solcher Einfluss wird *bedingte Wahrscheinlichkeit* genannt und mathematisch durch die Formel $P(A|B) = p$ ausgedrückt. Für das Beispiel wird angenommen, dass sich durch das Anlegen des magnetischen Feldes ($SP_{MF} = \{An, Aus\}$) die Wahrscheinlichkeit eine Augenzahl größer 3 zu würfeln verdreifacht. Für das Ergebnis $a_2 = 2$ beträgt die Wahrscheinlichkeit nun $P(a_2|b_{An}) = \frac{1}{12}$, wohingegen die Wahrscheinlichkeit für das Ergebnis $a_3 = 4$ nun auf $P(a_e|b_{An}) = \frac{1}{4}$ erhöht.

In einem Bayes'schen Netz kann ein solcher Zusammenhang als gerichtete Kante vom Knoten MF zum Knoten W visualisiert werden, allerdings ist auch die entgegengesetzte Richtung möglich. Die Austauschbarkeit der Richtung ist möglich, da durch eine Umformung des Satzes von Bayes (siehe Gleichung 1) die Berechnung aller Wahrscheinlichkeiten möglich ist.

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \quad (1)$$

In der vorliegenden Gleichung können die verschiedenen Komponenten wie folgt bezeichnet werden. $P(A)$ und $P(B)$ bezeichnen die sogenannten ursprünglichen Wahrscheinlichkeiten (A-priori-Wahrscheinlichkeiten) und $P(A|B)$ und $P(B|A)$ beschreiben die bedingten Wahrscheinlichkeiten, wobei $P(A|B)$ als A-posteriori-Wahrscheinlichkeit und $P(B|A)$ auch als Likelihood bezeichnet wird. Eine A-posteriori-Wahrscheinlichkeitsverteilung beschreibt die Auswirkungen von beobachteten Werten auf andere unbeobachtete Werte. Das Likelihood beschreibt eine Wahrscheinlichkeitsverteilung, die die möglichen Ursachen der beobachteten Werte beschreibt. Mit dem Satz von Bayes können somit nicht nur Auswirkungen berechnet werden, sondern auch Gründe für beobachtete Ereignisse. Dieser Umstand wird in Bayes'schen Netzen genutzt, um unter Unsicherheit zu schlussfolgern. Nachfolgend wird ein Beispiel vorgestellt, um den Sinn eines solchen Netzes und die Benutzung desselben zu verdeutlichen.

3.2.1.3 Beispiel

In Abbildung 13 ist ein von Charniak [34] entworfenes Bayes'sches Netz dargestellt. In dem Beispiel gilt es anhand der beobachtbaren Variablen zu bestimmen, ob die Familie im Haus ist oder nicht.

Die A-priori-Wahrscheinlichkeit, dass die Familie nicht im Haus ist ($SP_{FZ} = \{\text{Zuhause}, \text{Nicht Zuhause}\}$), beträgt 15 % ($P(FZ) = (0,85; 0,15)$). Für ein mögliches Darmproblem des Hundes ($SP_{DPH} = \{\text{ja}, \text{nein}\}$) besagt die A-priori-Wahrscheinlichkeitsverteilung, dass der Hund nur sehr selten dieses Problem hat ($P(DPH) = (0,99; 0,01)$).

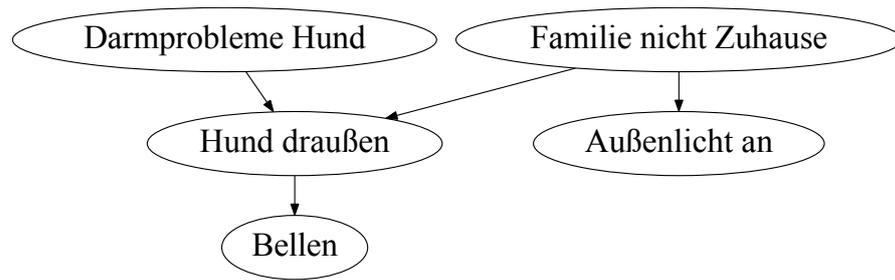


Abbildung 13: Beispiel für ein Bayes'sches Netz; Abbildung nach Charniak [34]

Familie ist	Zuhause	Nicht Zuhause
Licht an	0,05	0,6
Licht aus	0,95	0,4

Tabelle 16: Wahrscheinlichkeiten, dass das Außenlicht an bzw. aus ist

Familie ist	Zuhause		Nicht Zuhause	
Darmprobleme beim Hund	ja	nein	ja	nein
Hund draußen	0,97	0,3	0,99	0,9
Hund drinnen	0,03	0,7	0,01	0,1

Tabelle 17: Wahrscheinlichkeiten, dass der Hund drinnen oder draußen ist

Hund ist	Drinnen	Draußen
Bellen zu hören	0,01	0,7
Bellen nicht zu hören	0,99	0,3

Tabelle 18: Wahrscheinlichkeiten, dass das Bellen des Hundes zu hören ist

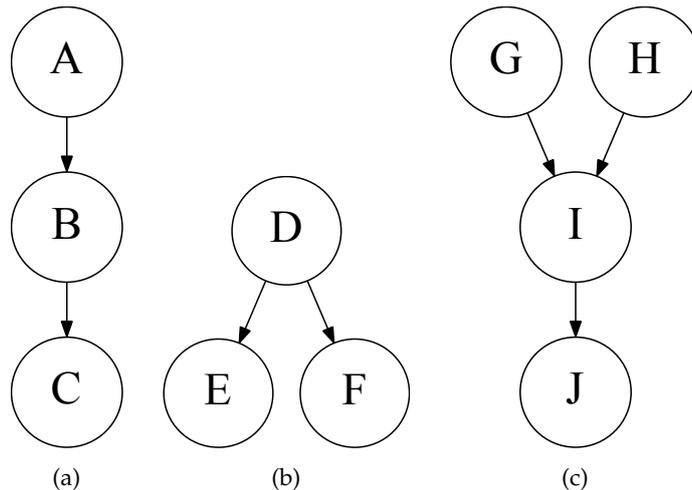


Abbildung 14: Beispiel für Verbindungen in Bayes'schen Netzen

Für die Knoten *Hund draußen*, *Außenlicht an* und *Bellen* müssen A-posteriori-Wahrscheinlichkeiten angegeben werden. Die entsprechenden Verteilungen sind in den Tabellen 16, 17 und 18 angegeben. Für Tabelle 16 beschreibt Charniak die bedingte Verteilung wie folgt: Wenn seine Frau das Haus verlässt, so schaltet sie oft das Licht an; allerdings schaltet Sie ebenfalls ab und zu das Licht an, wenn Sie im Haus ist und Gäste erwartet. Für den Aufenthaltsort des Hundes gilt (vgl. Tabelle 17), wenn die Frau von Charniak das Haus verlässt, so muss der Hund in den Garten. Der Hund muss ebenfalls in den Garten, wenn er ein Problem mit dem Darm hat. Weiter beschreibt Charniak, dass er den Hund meist nur bellen hören kann, wenn der Hund im Garten ist, allerdings gibt es in der Nachbarschaft andere Hunde, so dass er sich nicht sicher sein kann, ob das Bellen von seinem Hund kommt.

Um die Wahrscheinlichkeit zu berechnen, ob die Familie im Haus ist, werden Informationen über beobachtete Fakten (*Evidenz*) in das Netz eingegeben. Sind keine Informationen bekannt, so ändert sich nichts an der A-priori-Wahrscheinlichkeit; die Wahrscheinlichkeit, dass die Familie nicht im Haus ist, beträgt weiterhin 15 %.

Ist nur bekannt, dass der Hund ein Problem mit dem Darm hat, so ändern sich zwar die Wahrscheinlichkeiten, dass der Hund nicht im Haus ist (91 %) und dass das Bellen hörbar ist (64 %), es ändert sich allerdings nichts an der Einschätzung, ob die Familie im Haus ist (85 %) oder nicht (15 %). Der Grund für dieses Verhalten ist, dass die Knoten *Darmprobleme Hund* und *Familie nicht Zuhause* durch den Knoten *Hund draußen* d-separiert sind.

Die *d-Separation* beschreibt, wie sich Informationen durch das Netz weiterverbreiten bzw. an welchen Knoten die Informationen nicht weitergeleitet werden. In Abbildung 14 sind verschiedene Verbindungen zwischen Knoten dargestellt. Bei einer sequenziellen Verbindung,

wie sie in Abbildung 14a dargestellt ist, wird der Informationsfluss zwischen A und C unterbrochen, wenn der Zustand von B bekannt ist. Um es mit dem Beispiel aus der Abbildung 13 zu erklären; wenn bekannt ist, dass der Hund im Haus ist, dann hat die zusätzliche Information, dass die Familie im Haus ist, keinen Einfluss darauf, ob das Bellen des Hundes zu hören ist oder nicht. Bei einer divergierenden Verbindung, wie sie in Abbildung 14b dargestellt ist, fließt die Information von E nach F nur, wenn über D nichts bekannt ist. Am Beispiel erklärt bedeutet das, dass der Umstand, dass der Hund draußen ist, nichts an der Erkenntnis über das Außenlicht ändert, wenn bekannt ist, dass die Familie im Haus ist (oder nicht ist). Bei einer konvergierenden Verbindung, wie sie in Abbildung 14c zu sehen ist, ändert sich allerdings das bisher bekannte Verhalten. Hier können nur Informationen zwischen G und H fließen, wenn der Zustand in einem der Knoten I, J oder einem beliebigen Kinderknoten von I bekannt ist. Für das Beispiel bedeutet das; wenn unsicher ist, wo der Hund ist und unsicher ist ob das Bellen vom Hund der Familie kommt, so ändert das Wissen über den Aufenthaltsort der Familie nichts am Erkenntnisstand über die Darmprobleme des Hundes. Ist allerdings die Information gegeben, dass das Bellen vom Familienhund ist oder dass der Hund im Garten ist, so beeinflussen sich die beiden Knoten *Darmprobleme Hund* und *Familie nicht Zuhause* gegenseitig.

3.2.1.4 Anwendung

Bayes'sche Netze werden in unterschiedlichen Bereichen und für unterschiedliche Zwecke eingesetzt. Nachfolgenden werden die vier häufigsten Zwecke kurz vorgestellt und jeweils an einem Beispiel illustriert.

Der erste Einsatzzweck eines Bayes'schen Netzes ist die Berechnung der Wahrscheinlichkeit, dass ein definiertes Ereignis eintritt (*Probability of Evidence; $P(e)$*). Als Beispiel wird das Netz aus dem vorhergehenden Abschnitt 3.2.1.3 verwendet. Die Frage lautet: Wie hoch ist die Wahrscheinlichkeit, dass das Außenlicht ausgeschaltet und der Hund im Haus ist? Antwort: Die Probability of Evidence beträgt $P(e) = 0,566$. In 56,6 % der Fälle kann somit beobachtet werden, dass das Außenlicht ausgeschaltet und der Hund im Haus ist.

Der zweite Einsatzzweck dient dazu, die Wahrscheinlichkeitsverteilungen vor und nach dem Einfügen von Evidenz zu vergleichen (*Prior Marginals und Posterior Marginals*), um Auswirkungen von Ereignissen bewerten zu können. Für das Beispiel wird der Knoten *Familie nicht Zuhause* aus der Abbildung 13 untersucht. Bevor Evidenz in das Netz eingefügt wird, beträgt die Wahrscheinlichkeit, dass die Familie nicht Zuhause ist, 15 %. Wird nun festgestellt, dass das Außenlicht aus und der Hund im Haus ist, so sinkt die Wahrscheinlichkeit, dass die Familie nicht im Haus ist auf knapp 1 %. Mit der vorliegenden Evidenz ist die Wahrscheinlichkeit, dass die Familie Zuhause ist, somit sehr hoch.

Im dritten Fall wird ein Bayes'sches Netz verwendet, um die Ausprägung der Knoten zu finden, die die beobachteten Verhältnisse (Evidenz) am besten erklären (*Most Probable Explanation*). Für das schon verwendete Beispiel (Außenlicht aus; Hund im Haus) lautet die wahrscheinlichste Belegung der nicht beobachteten Knoten wie folgt: Die Familie ist Zuhause, der Hund hat kein Problem mit seinem Darm und es ist kein Bellen zu hören.

Im vierten und letzten Fall, wird ein Bayes'sches Netz verwendet, um die Wahrscheinlichkeiten von verschiedenen Hypothesen zu untersuchen bzw. die Hypothese auszuwählen, die am wahrscheinlichsten ist (*Maximum A Posteriori Hypothesis*). Wird z. B. beobachtet, dass der Hund im Garten ist, so existieren hierfür verschiedene Hypothesen, die die Beobachtung erklären. Mit einem Bayes'schen Netz kann berechnet werden, welche Hypothese die höchste Wahrscheinlichkeit hat.

3.2.1.5 Vor- und Nachteile von Bayes'schen Netzen bei der Modellierung von Nutzerattributen

Um Nutzerattribute adäquat zu repräsentieren und demografische Nutzermodelle zu generieren müssen demografische Verteilungen und deren Zusammenhänge abgebildet werden können. Ein Vorteil eines Bayes'schen Netzes ist, dass für die Ausprägungen eines Knotens eine Verteilung angegeben werden kann, die die Verteilung für das entsprechende Attribut in der Bevölkerung widerspiegelt. Der zweite Vorteil ist, dass durch Kanten zwischen Knoten die Beziehungen zwischen Variablen abgebildet werden können, was Wissen über Zusammenhänge modellierbar macht. Der letzte, und entscheidende, Vorteil ist, dass aufgrund von wenigen beobachteten Merkmalen, auf die Ausprägung bzw. die Verteilung von Ausprägungen von anderen Knoten geschlossen werden kann. Durch diesen Umstand ist es möglich, dass ein späterer Nutzer der MeMo-Werkbank valide demografische Nutzermodelle erzeugt, ohne über das entsprechende Expertenwissen zu verfügen.

Ein Nachteil der aus der Verwendung von Bayes'schen Netzen entsteht ist, dass in einem Netz die Verteilung für die Ausprägung in einem Knoten für jede Kombination der Elternknoten angegeben werden muss. Hat ein Knoten viele Elternknoten und diese wiederum viele Ausprägungen, so ist die resultierende Wahrscheinlichkeitstabelle für den zu modellierenden Knoten sehr groß. Weiterhin besteht die Notwendigkeit für alle Kombinationen eine Verteilung angeben zu müssen, was nicht immer möglich ist.

Werden die Vor- und Nachteile gegeneinander abgewogen, so zeigt sich, dass die durch die Flexibilität der Lösung und durch die Validität der erzielbaren Ergebnisse gewonnenen Vorteile schwerer wiegen, als die genannten Nachteile.

3.2.2 Repräsentation der Nutzerattribute

Für die Umsetzung des Bayes'schen Netzes in Software, wurde die Implementierung von Druzdzel (GeNIe/SMILE) [56] verwendet. Da die Lizenz der Software inkompatibel mit der Implementierung der MeMo-Werkbank ist, kann das Bayes'sche Netz nicht direkt in die Werkbank integriert werden. Als Alternative zur direkten Einbindung wurde ein Client-Server-Ansatz gewählt. Beim Client-Server-Ansatz fragt die MeMo-Werkbank über ein Application Programming Interface (API) einen Server an. Auf dem Server werden die Nutzermodelle aus dem Netz erzeugt, in eine JavaScript Object Notation (JSON) überführt und der Werkbank zur Verfügung gestellt. In der MeMo-Werkbank werden die so erhaltenen Nutzermodellbeschreibungen in Nutzermodelle umgewandelt und für die Simulation verwendet. Auf eine detaillierte Beschreibung der Implementierung des Bayes'schen Netzes wird an dieser Stelle verzichtet, es soll nur beschrieben werden, wie der Abruf von Nutzermodellbeschreibungen funktioniert.

Wird mit der MeMo-Werkbank die Interaktion mit einem Systemmodell simuliert, so muss vom Modellierer mindestens eine Nutzermodellgruppe ausgewählt werden, die für die Simulation verwendet wird. Wie in Abschnitt 2.1.5 ausgeführt, beschreibt eine Nutzermodellgruppe verschiedene Eigenschaften bzw. Einschränkungen von zugeordneten Nutzermodellen. Bei der Verwendung des Bayes'schen Netzes zur Generierung von Nutzermodellen, werden die in der Nutzergruppe definierten Eigenschaften und Einschränkungen als Evidenz (vgl. Abschnitt 3.2.1.3) ins Netz eingebracht. Hiernach werden durch einen Algorithmus, entsprechend der Anzahl der Iterationen in der Simulation, die Nutzermodelle aus dem Netz generiert. Sind alle Nutzermodelle erzeugt, so werden diese in eine JSON-Repräsentation überführt und an die MeMo-Werkbank übertragen. In der Werkbank werden die Nutzermodellbeschreibungen in MeMo-Nutzermodelle übertragen und in der Simulation verwendet.

3.2.2.1 Übersicht der Nutzerattribute

In den Tabellen 19 und 20 sind verschiedenen Nutzerattributen und deren Zusammenhänge bzw. Auswirkungen dargestellt. Die Angaben in den Tabellen leiten sich aus den Erkenntnissen ab, die im Abschnitt 3.1 beschrieben wurden. Von den insgesamt 24 Attributen werden 8 Attribute (Cholesterinspiegel, Einkommen, Ernährung, Innenohrinfektionen, Lärm, Medikamente, psychische Erkrankungen und Stress) nicht in das Bayes'sche Netz übernommen. Die Attribute *Cholesterinspiegel*, *Innenohrinfektionen* und *psychische Erkrankungen* werden aufgrund fehlender oder unzureichender Daten ausgeschlossen. Das Attribut *Einkommen* geht in dem Attribut *Soziale Schicht* auf und die Attribute *Lärm*, *Ernährung*, *Medikamente* und *Stress* sind so unspezifisch, dass eine Modellierung in der Form nicht sinnvoll erscheint.

Attribute	Auswirkungen
Adipositas	Demenz, Arthrose, Gicht, Diabetes mellitus, Herz-Kreislauf-Erkrankungen (Herzinfarkt), Bluthochdruck, Bewegungsapparat
Alkoholkonsum	Demenz, Schlaganfall, Haltetremor, Gicht, Bluthochdruck
Alter	Hören, Seheinschränkungen, Grauer Star, altersabhängige Makuladegeneration, fluide Intelligenz, kristalline Intelligenz, Demenz, Schlaganfall, Schädel-Hirn-Trauma, essenzieller Tremor, Arthrose, rheumatoide Arthritis, Gicht, Schwerbehinderung, Bluthochdruck, Diabetes mellitus, Adipositas, Unfall
Bildung	Demenz, Arthrose
Bluthochdruck	Hören, Demenz, Schlaganfall, Seheinschränkungen
Cholesterinspiegel	Schlaganfall
Depression	Denk- und Konzentrationsstörung, verlangsamte Reaktion, Bewegungsstörungen
Diabetes mellitus	Hören, Seheinschränkungen, Demenz, Schlaganfall, Gicht, Herz-Kreislauf-Erkrankungen (Herzinfarkt), Nervenerkrankungen, Amputationen
Einkommen	Hören
Ernährung	Bluthochdruck, Diabetes mellitus, Adipositas
Genetische Faktoren	Hören, Kurzsichtigkeit, Weitsichtigkeit, Demenz, Tremor, Arthrose, rheumatoide Arthritis, Diabetes mellitus, Adipositas, Depression, Multiple Sklerose

Tabelle 19: Nutzerattribute und deren Auswirkungen – Teil 1

Attribute	Auswirkungen
Geschlecht	Hören, Seheinschränkungen , Intelligenzminderung, Schlaganfall, Schädel-Hirn-Trauma, Arthrose, Gicht, temporäre Einschränkungen, Bluthochdruck, Multiple Sklerose, Depression, Unfall
Herzrhythmusstörungen	Schlaganfall
Hören	Depression
Innenohrinfektionen	Hören
Lärm	Hören
Medikamente	Hören
Multiple Sklerose	Sensibilitätsstörungen, Seheinschränkungen, Lähmungen
Psychische Erkrankungen	Demenz
Rauchen	Hören, altersabhängige Makuladegeneration, Schlaganfall, rheumatoide Arthritis, Diabetes mellitus
Soziale Schicht	Hören, Intelligenzminderung, Schlaganfall, Depression
Stress	Demenz, Bluthochdruck, Diabetes mellitus
Unfall	Körperliche Einschränkung
Wenig Aktivitäten	Demenz, Schlaganfall, Bluthochdruck, Diabetes mellitus, Adipositas

Tabelle 20: Nutzerattribute und deren Auswirkungen – Teil 2

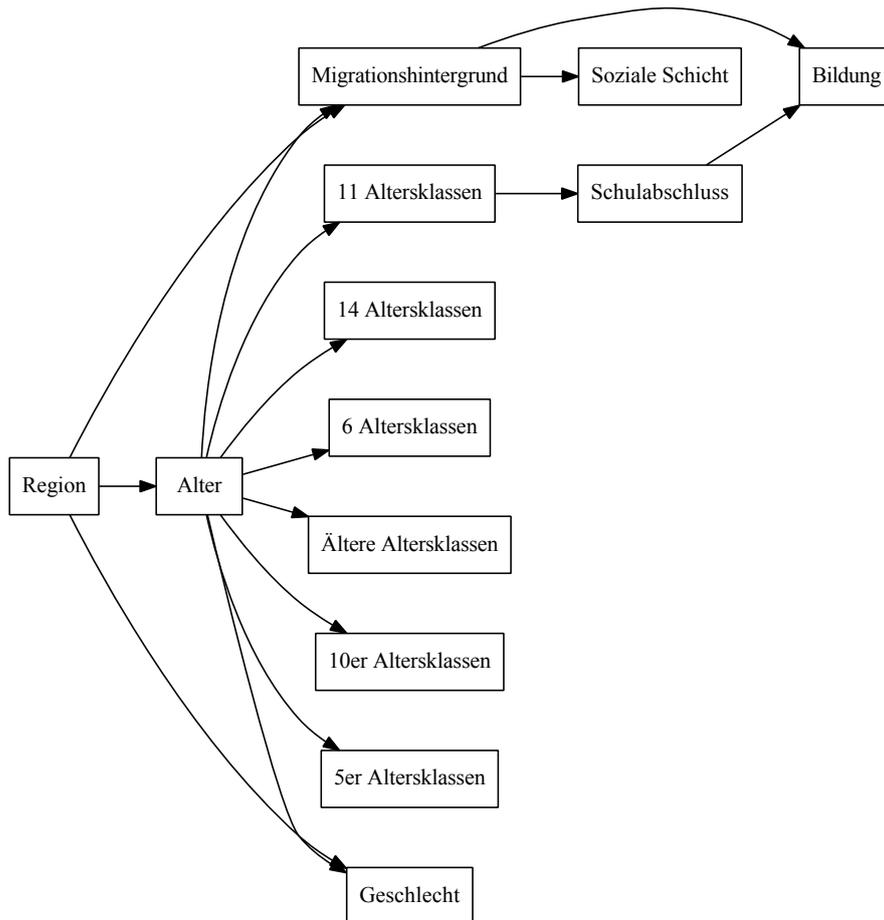


Abbildung 15: Bayes'sches Netz: Zusammenhänge der allgemeinen Attribute (Teilnetz)

Zusätzlich zu den in den Tabellen 19 und 20 aufgeführten Attributen werden im Bayes'schen Netz noch Angaben zur *Region*, zum *Migrationshintergrund*, bezüglich der *Sprachkenntnisse*, *Technikaffinität* und *Technikeinstellung* aufgenommen. Für die Begründungen, warum diese Knoten in das Netz aufgenommen werden müssen, wird auf die nachfolgenden Abschnitte verwiesen.

Da das entstandene Netz sehr groß und in der Abbildung sehr unübersichtlich ist, wird auf die Gesamtdarstellung verzichtet. In den nachfolgenden Unterabschnitten werden deshalb Teilnetze vorgestellt. Sofern ein Knoten von einem Knoten aus einem anderen Teilnetz abhängt, wird der Elternknoten in die Darstellung des Teilnetzes aufgenommen und grau dargestellt.

3.2.2.2 Allgemeine Attribute

Unter *allgemeine Attribute* sind Angaben zur geografischen Herkunft (*Region*), zum *Migrationshintergrund*, zur Alters- & Geschlechterverteilung, zur sozialen Schicht, dem Bildungsgrad und zu den verschie-

denen Altersgruppierungen zusammengefasst. In Abbildung 15 ist dargestellt, wie diese Attribute zusammenhängen.

Für das vorliegende Teilnetz hat der Knoten *Region* 16 Ausprägungen, die die Bundesländer in Deutschland abdecken. Der Knoten *Region* wurde in das Netz aufgenommen, da die Alters- und Geschlechterverteilung in den einzelnen Bundesländern unterschiedlich ist, weiterhin ist der Anteil der Personen mit Migrationshintergrund in den Regionen sehr ungleich.

Der Knoten *Alter* enthält 101 verschiedene Ausprägungen und ermöglicht die Abbildung der verschiedenen Altersstufen von *unter 1 Jahr* bis *100 Jahre und älter*. Die Daten für die prozentuale Verteilung pro Altersstufe sind der Datenbank des Zensus 2011 [221] entnommen. Der Knoten *Geschlecht*, hat zwei Ausprägungen und ist vom Knoten *Alter* und *Bundesland* abhängig, da der Frauenanteil im höheren Alter steigt und in den einzelnen Regionen verschieden hoch ist. Die Daten, die die Verteilungen für den Knoten *Geschlecht* beschreiben, sind ebenfalls der Datenbank des Zensus 2011 entnommen.

Zusätzlich zum Knoten *Alter* sind noch verschiedene Hilfsknoten (*11 Altersgruppen*, *14 Altersgruppen*, *6 Altersgruppen*, *Ältere Altersgruppe*, *10er Altersgruppen* und *5er Altersgruppen*) modelliert. Die verschiedenen Knoten mit unterschiedlichen Altersgruppen werden benötigt, da für verschiedene Einschränkungen (Schwerhörigkeit, Sehbeeinträchtigungen, Demenzerkrankungen, usw.) nur Angaben für Altersgruppen vorliegen (siehe Abschnitt 3.1).

Bei den Knoten *Soziale Schicht* und *Schulabschluss* wird eine Abhängigkeit vom Migrationshintergrund modelliert. Durch dieses Vorgehen wird zum einen eine zirkuläre Abhängigkeiten zwischen der Einordnung in eine soziale Schicht und dem Bildungserfolg [172, 218] vermieden, zum anderen sind beide Attribute stark vom Migrationshintergrund beeinflusst [75, S. 299 ff.]. Die Verteilung für den binären Knoten *Migrationshintergrund* wird in Abhängigkeit vom Alter und Region definiert, da der Anteil der Personen mit Migrationshintergrund zwischen den Bundesländern sehr verschiedenen ist und in den höheren Lebensaltern stark abnimmt. Der Knoten *Soziale Schicht*, der die Verteilung des sozialen Status repräsentiert, hat drei Ausprägungen *Unterschicht*, *Mittelschicht* und *Oberschicht*.

Der Knoten *Schulabschluss* hat insgesamt 8 Ausprägungen (*Nicht anwendbar*, *Ohne Schulabschluss*, *noch Schüler*, *Hauptschulabschluss*, *Realschulabschluss*, *Schüler der Oberstufe*, *Fachhochschulreife*, *Abitur*). Die Ausprägung *Nicht anwendbar*, wird nur für Personen unter 15 Jahren verwendet, da für diese eine Schulpflicht besteht und in der Regel kein Schulabschluss vorliegt. Die Ausprägungen *noch Schüler* und *Schüler der Oberstufe* werden hauptsächlich für Personen in den Altersgruppen *15 bis 17* und *18 bis 24* (Knoten: *11 Altersklassen*) verwendet, da in diesen Altersgruppen ein Teil der Bevölkerung noch zur Schule geht,

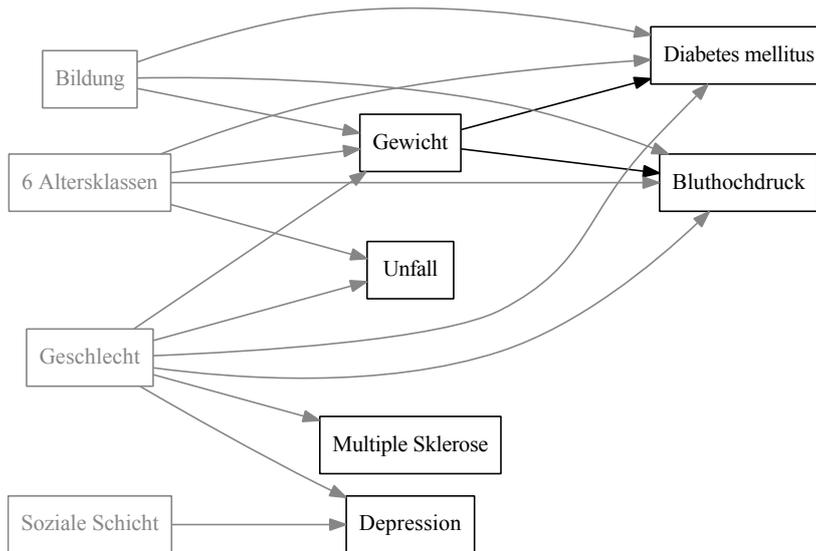


Abbildung 16: Bayes'sches Netz: Erkrankungen und Unfälle (Teilnetz)

andere Teile demgegenüber eine Ausbildung absolvieren oder einem Beruf nachgehen.

Die Verteilung der Wahrscheinlichkeiten im Knoten *Bildung* hängt von der Ausprägung des *Schulabschluss*-Knotens ab. Ist der Knoten *Schulabschluss* mit *nicht anwendbar* oder mit *noch Schüler* markiert, so wird der Knoten *Bildung* mit *nicht anwendbar* markiert, da keine Aussage über das Bildungsniveau getroffen werden kann. Die Zuordnung der anderen Zustände des Bildungsknotens (*niedrig*, *mittel*, *hoch*) erfolgt anhand des Schulabschlusses.

3.2.2.3 Erkrankungen und Unfälle

In diesem Teil der Arbeit wird der Ausschnitt aus dem Netz vorgestellt, der die in Abschnitt 3.1 genannten Erkrankungen und Unfälle enthält. Die Abbildung 16 visualisiert die Zusammenhänge zwischen den Attributen. Mit Grau sind die Knoten dargestellt, die aus anderen Teilnetzen stammen.

Der Knoten *Gewicht* hat vier Ausprägungen (*Untergewicht*, *Normalgewicht*, *Übergewicht* und *Adipositas*). Die Daten über die Verteilung in der Bevölkerung stammen aus dem Bericht von Mensik et al. [161] und aus der Datenbank des GBE [224]. Abhängig vom Geschlecht und Bildungsniveau steigt mit zunehmendem Alter die Prävalenz für Übergewicht und Adipositas, was die Abhängigkeit von den Knoten *6 Altersklassen*, *Geschlecht* und *Bildung* erklärt.

Der Knoten *Bluthochdruck* hat vier Elternknoten (*Gewicht*, *6 Altersklassen*, *Geschlecht* und *Bildung*) und eine binäre Ausprägung (*liegt vor* und *liegt nicht vor*). Die zur Beschreibung der Verteilung verwendeten Daten entstammen dem Gesundheitsbericht des Bundes – Heft 43 (Hy-

pertonie) [114] und der Datenbank zur Gesundheitsberichtserstattung des Bundes (GBE) [224].

Eine Depression betrifft Frauen öfter als Männer und Angehörige aus den unteren sozialen Schichten häufiger als Personen aus höheren sozialen Schichten, deshalb wurden Kanten von den Knoten *Geschlecht* und *Soziale Schicht* zum Knoten *Depression* gezogen. Eine Depression hat verschiedene Schweregrade, allerdings wird im vorliegenden Netz nur zwischen drei Ausprägungen (*keine Depression*, *depressive Symptomatik* und *Depression*) unterschieden. Die Daten bezüglich der Prävalenz von Depressionen in der deutschen Bevölkerung entstammen der Studie und Busch et al. [25] und für die jüngeren Personen (unter 40 Jahren) aus der GBE Datenbank [224].

Für den *Diabetes mellitus*-Knoten ist eine Abhängigkeit von den 6 Altersklassen, dem Geschlecht, dem Gewicht und dem Bildungsniveau modelliert. Beim Diabetes wird nicht zwischen dem Typ-1-Diabetes und dem Typ-2-Diabetes unterschieden, weshalb der Knoten eine binäre Ausprägung (*liegt vor* und *liegt nicht vor*) hat. Die Daten über die Verteilung von Diabetes mellitus in der Bevölkerung sind der GBE Datenbank [224] und drei Studien [96, 97, 193] entnommen.

Wie bereits in Abschnitt 3.1.5 erwähnt, sind die ursächlichen Auslöser für eine Multiple Sklerose (MS) weitestgehend unbekannt. Frauen erkranken häufiger als Männer an MS und die Verbreitung in den nördlichen Industrieländern ist höher als in den Ländern, die näher am Äquator liegen. Um die Verbreitung zu modellieren wird auf die Arbeiten von Yaldizli & Kappos [256] und Hoffmann et al. [100] zurückgegriffen. Die Ausprägung des Knotens ist binär (*liegt vor* und *liegt nicht vor*).

Um die Verbreitung von Unfälle in der Bevölkerung zu modellieren werden Daten des Robert-Koch-Institutes [237], der Deutschen Gesetzlichen Unfallversicherung [220] und vom Gesundheitssurvey 2004 [203] verwendet. Die Abhängigkeit von Alter und Geschlecht resultiert aus dem Umstand, dass insbesondere jüngere Männer von nicht-tödlichen Unfällen betroffen sind, die eine ärztliche Behandlung notwendig machen (vgl. Abschnitt 3.1.4).

3.2.2.4 *Einschränkungen in der Wahrnehmung*

In diesem Abschnitt werden die Verteilungen für die Einschränkungen im Hören und Sehen – die Wahrnehmung betreffende Attribute – modelliert.

Bei den Höreinschränkungen wird zwischen *keiner Einschränkung*, einer *leichten* bzw. *schweren Schwerhörigkeit* und *Taubheit* unterschieden. Die Verteilung im Knoten wird durch das Alter, das Geschlecht und durch das Bildungsniveau (Lärmbelastung im Beruf) beeinflusst (siehe Abschnitt 3.1.1.1). Als Risikofaktor für eine Höreinschränkung gilt weiterhin das Vorliegen von Diabetes mellitus. Als Datenquellen zur

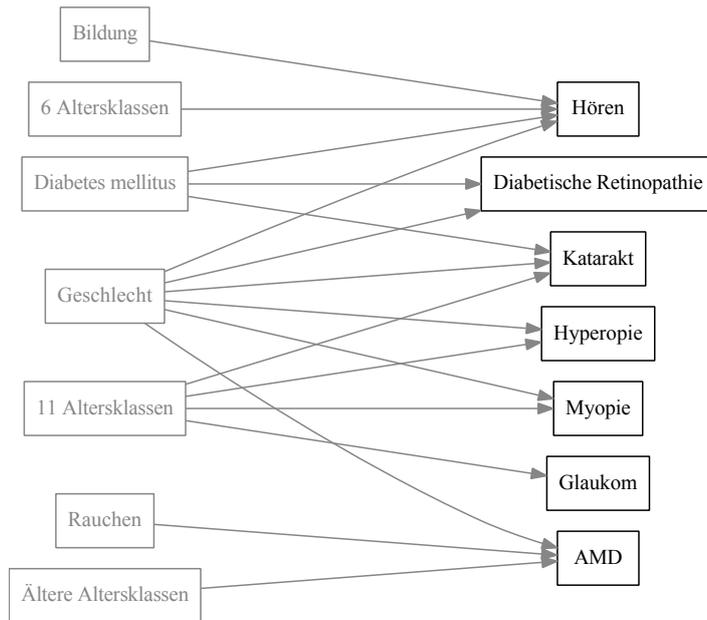


Abbildung 17: Bayes'sches Netz: Einschränkungen in der Wahrnehmung (Teilnetz)

Modellierung der Verteilung in der Gesellschaft werden verschiedene Studien [85, 95, 257] und Bücher [16, 152, 249] verwendet.

Bei der visuellen Wahrnehmung wird hinsichtlich der Art der Einschränkung unterschieden. Im vorliegenden Netz werden 6 verschiedene Knoten modelliert, die den Visus einschränkende Attribute beschreiben; hierbei handelt es sich um *diabetische Retinopathie*, *Myopie* (Kurzsichtigkeit), *Hyperopie* (Weitsichtigkeit), *Katarakt*, *altersabhängige Makuladegeneration* und *Glaukom*.

Bei der diabetischen Retinopathie wird zwischen drei Zuständen (*keine*, *leichte* und *schwere Retinopathie*) unterschieden. Wie in Abschnitt 3.1.2 dargestellt, sind Männer häufiger als Frauen betroffen, weshalb die Kante vom Knoten *Geschlecht* gezogen wurde. Weiterhin ist das Vorliegen einer Diabetes-mellitus-Erkrankung die Voraussetzung für eine diabetische Retinopathie.

Bei den Knoten *Myopie* und *Hyperopie* werden jeweils drei Ausprägungen (*schwer*, *leicht* und *keine Einschränkung*) modelliert. Da die Einschränkungen in der Fehlsichtigkeit mit dem Alter zunehmen und Frauen stärker als Männer betroffen sind, werden Kanten von den Knoten *11 Altersklassen* und *Geschlecht* eingefügt.

Der Knoten *Glaukom* wird binär modelliert, da nicht zwischen den verschiedenen Glaukom-Formen (Offenwinkelglaukom > 90%) unterschieden wird. In jungen Jahren sind nur wenige Personen betroffen (1:10.000) [232, S. 120 ff.], allerdings steigt die Zahl der Betroffenen mit dem Alter stark an, so sind über 7 % der Personen über 75 Jahren betroffen.

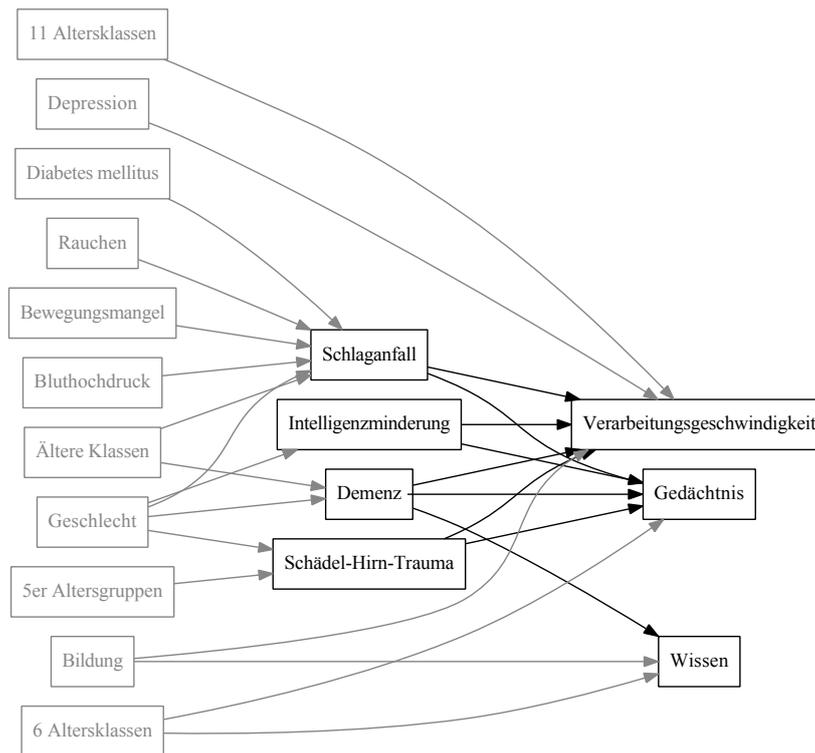


Abbildung 18: Bayes'sches Netz: Einschränkungen in der Kognition (Teilnetz)

Beim Katarakt wird zwischen *liegt nicht vor* und *liegt vor* unterschieden. Als beeinflussende Faktoren für die Entwicklung eines Katarakts werden Abhängigkeiten vom Alter, vom Geschlecht und das Vorliegen einer Diabetes-Erkrankung im Netz modelliert.

Bei der altersabhängige Makuladegeneration (AMD) wird der Knoten ebenfalls mit zwei Ausprägungen modelliert. Da insbesondere ältere Menschen betroffen sind, wird die Abhängigkeit von den älteren Altersklassen definiert. Die Kanten von den Knoten *Geschlecht* und *Rauchen* reflektieren das höhere Risiko von Frauen und von Rauchern an einer AMD zu erkranken. Die Daten für die Verteilung der Einschränkungen in der Gesellschaft für das Katarakt und die AMD sind verschiedenen Studien [39, 38, 136, 174], Büchern [134, 146, 147, 151] und der GBE Datenbank [224] entnommen.

3.2.2.5 Einschränkungen in der Kognition

Bei den kognitiven Fähigkeiten werden allgemeine und einschränkende Faktoren modelliert (siehe Abbildung 18). Zu den allgemeinen Faktoren zählen die fluide Intelligenz (*Gedächtnis* und *Verarbeitungsgeschwindigkeit*) und die kristalline Intelligenz (*Wissen*). Bei den einschränkenden Faktoren werden Knoten wie *Demenz*, *Intelligenzminderung*, *Schlaganfall* und *Schädel-Hirn-Trauma* modelliert. Wie der

Abbildung zu entnehmen ist, sind die Ursachen für Erkrankungen vielfältig und das Zusammenspiel ist komplex.

Die fluide Intelligenz wird mit Netz durch die Knoten *Verarbeitungsgeschwindigkeit* und *Gedächtnis* repräsentiert, wobei der Knoten *Gedächtnis* das Arbeits- und Kurzzeitgedächtnis repräsentiert. Der Knoten *Verarbeitungsgeschwindigkeit* hat drei Ausprägungen *normal*, *verlangsamt* und *nicht anwendbar* und ist in Abhängigkeit von dem Alter, der Bildung, Schlaganfall, Schädel-Hirn-Trauma und dem Vorliegen einer Intelligenzminderung oder Depression modelliert. Die Ausprägung *nicht anwendbar* wird verwendet, wenn die der Knoten *Intelligenzminderung* die Ausprägungen *mittel*, *schwer* oder *schwerste* annimmt.

Die Ausprägung des *Gedächtnis*-Knotens ist binär (*normal* oder *eingeschränkt*) modelliert. Wie bei dem Knoten *Verarbeitungsgeschwindigkeit* werden Abhängigkeiten von den Knoten *Intelligenzminderung*, *Schlaganfall*, *Schädel-Hirn-Trauma* und *Alter* hinzugefügt. Zusätzlich wird eine Kante vom Knoten *Demenz* gezogen, um die Einschränkungen einer Demenz abzubilden.

Die kristalline Intelligenz wird durch den Knoten *Wissen* dargestellt. In diesem Knoten wird die Ausprägung durch eine Dreiteilung in *niedrig*, *mittel* und *hoch* beschrieben. Wie der Abbildung zu entnehmen ist, wird die kristalline Intelligenz vom Alter, der Bildung und dem Vorliegen einer Demenz-Erkrankung beeinflusst.

Eine Intelligenzminderung kann in 5 Stufen (*keine Einschränkung*, *leichte*, *mittelschwere*, *schwere* und *schwerste Intelligenzminderung*) beschrieben werden (siehe Abschnitt 3.1.3.4). Da Männer häufiger als Frauen betroffen sind, ist der entsprechende Knoten *Geschlecht* durch eine Kante verbunden.

Von einem Schädel-Hirn-Trauma (SHT) sind jüngere Männer stärker als andere Personengruppen betroffen, weshalb von den Knoten *5er Altersgruppen* und *Geschlecht* jeweils eine Kante zum Knoten *Schädel-Hirn-Trauma* führt. Ein SHT kann verschiedene Schweregrade haben, weshalb der Knoten vier Ausprägungen (*kein*, *leichtes*, *mittleres* und *schwerwiegendes Trauma*) hat.

Bei der Demenz kann die Eingliederung anhand des Stadiums der Erkrankung erfolgen. Im Allgemeinen wird zwischen einem frühen, mittleren und späten Stadium unterschieden. Im Netz wird der Knoten allerdings binär (*Ja* oder *Nein*) modelliert, da entsprechende Daten über Verteilungen in den Stadien fehlen. Die Prävalenz von Demenz-Erkrankungen hat eine starke Abhängigkeit vom Alter. Weiterhin sind mehr Frauen als Männer von der Krankheit betroffen.

Der Schlaganfall ist, neben dem SHT, die häufigste Krankheit, die zu einer Hirnschädigung führt. Wie in Abbildung 18 dargestellt, sind die Ursachen und Risiken für einen Schlaganfall sehr unterschiedlich. Als empirische gesicherte Faktoren gelten unter anderem das chronologische Alter, männliches Geschlecht, Bluthochdruck, Diabetes mellitus

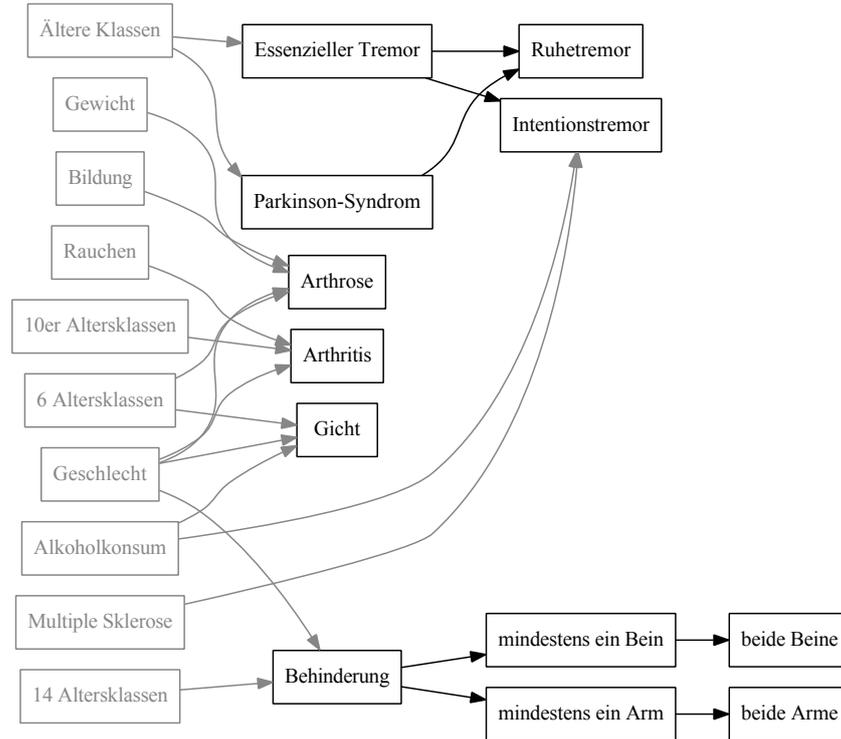


Abbildung 19: Bayes'sches Netz: Einschränkungen in der Motorik (Teilnetz)

und Rauchen. Ein Mangel an Bewegung wird allerdings ebenfalls als Risikofaktor gesehen [189, S. 53 ff]. Im Netz wird der Knoten binär (*erlitten* oder *nicht erlitten*) modelliert.

Zur Modellierung der Verteilungen von kognitiven Einschränkungen wird auf verschiedene Studien [72, 181, 191, 227] und Bücher [14, 179, 189] zurückgegriffen. Weiterhin werden Daten aus der Datenbank zur Gesundheitsberichtserstattung des Bundes [224] verwendet.

3.2.2.6 Einschränkungen in der Motorik

Im Teilnetz, das die Einschränkungen in der Motorik abbildet, werden die Knoten für verschiedene Tremor-Arten (Essenzieller Tremor, Ruhetremor und Intentionstremor), rheumatischen Erkrankungen (Arthrose, Arthritis und Gicht), Körperbehinderungen (bezüglich der Beine und Arme) und das Parkinson-Syndrom aufgenommen. In Abbildung 19 sind die Beziehungen zwischen den Knoten dargestellt.

Als ursächliche Risikofaktoren für die Entwicklung einer Arthrose sind das Alter, das Geschlecht, das Gewicht und ein niedriger Bildungsstand (bei Frauen) bekannt, weshalb die entsprechenden Knoten mit dem Knoten *Arthrose* verbunden sind. Die Ausprägung für den Knoten ist binär (*liegt vor* und *liegt nicht vor*).

Hinsichtlich der Ausprägung und Auswirkungen entsprechen die Knoten *Gicht* und *Arthritis* dem *Arthrose*-Knoten. Die Prävalenz von

Gicht wird durch unterschiedliche Faktoren (Alkoholkonsum, Alter und Geschlecht) beeinflusst, weshalb die entsprechenden Knoten über Kanten mit dem Knoten *Gicht* verbunden sind. Der Knoten *Arthritis* hat, aufgrund der Zusammenhänge, die Knoten *Altersgruppe*, *Geschlecht* und *Rauchen* als Elternknoten.

Die Ausprägungen der Knoten *14 Altersklassen* und *Geschlecht* beeinflussen die Wahrscheinlichkeit für das Auftreten einer Körperbehinderung, weshalb der binäre Knoten *Körperbehinderung* durch eine Kante mit diesem verbunden ist. Da auch Körperbehinderungen unterschiedliche Häufigkeiten haben, werden die 4 Knoten *mindestens ein Bein*, *mindestens ein Arm*, *beide Beine* und *beide Arme* hinzugefügt, die jeweils eine binäre Ausprägung haben.

Wie in Abschnitt 3.1.4 erläutert, existieren verschiedene Tremor-Arten, deren Prävalenz durch unterschiedliche Ursachen (siehe Abbildung 19) beeinflusst wird und deren Auftreten die Wahrscheinlichkeit für andere Tremor-Arten beeinflusst. Wie in Abschnitt 3.1.4.2 beschrieben, sind die Ursachen für eine Tremor-Erkrankung oftmals unklar, allerdings treten bei Personen mit bestimmten Eigenschaften, diese Erkrankungen gehäuft auf. Diesem Umstand wird, durch die Kanten, Rechnung getragen. Da keine verlässlichen Daten über die Verteilung der Schweregrade vorliegen, werden die Knoten mit binären Ausprägungen modelliert.

Der letzte Knoten in diesem Teilnetz bildet die Verteilung des Parkinson-Syndroms in der Bevölkerung ab. Wie bei den Knoten zur Beschreibung von Tremor-Erkrankungen wird eine binäre Ausprägung modelliert. Da ältere Menschen häufiger als jüngere von Parkinson betroffen sind, wird eine Kante vom Knoten *Ältere Klassen* zum Knoten *Parkinson-Syndrom* eingefügt.

Als Datenbasis zur Modellierung der Verteilungen werden verschiedene Bücher [171, 185], Studien [61, 197, 230] und Berichte [183, 220, 224, 237] genutzt.

3.2.2.7 Technikaffinität und Technikeinstellung

Anders als bei den vorhergehenden Abschnitten, werden in diesem Abschnitt keine Einschränkungen oder Erkrankungen abgebildet, sondern Eigenschaften und Einstellungen einer Person (*Technikaffinität* und *Technikeinstellung*) modelliert.

Das Attribut *Technikaffinität* beschreibt ein Merkmal einer Person, dass die Anziehung eines Menschen zu technischen Geräten, deren positive Einstellung, Begeisterung und Vertrauen in Technik ausdrückt [127]. Wie dieser Definition zu entnehmen ist, handelt es sich bei der Technikaffinität um kein eindimensionales Konstrukt, weshalb Technikaffinität anhand von 4 Dimensionen (Begeisterung für Technik, positive bzw. negative Folgen von Technikeinsatz und Kompetenz im Umgang mit Technik) beschrieben wird [127]. Zur Erfassung der Attribute kann der TA-EG Fragebogen von Karrer et al. [127] verwen-

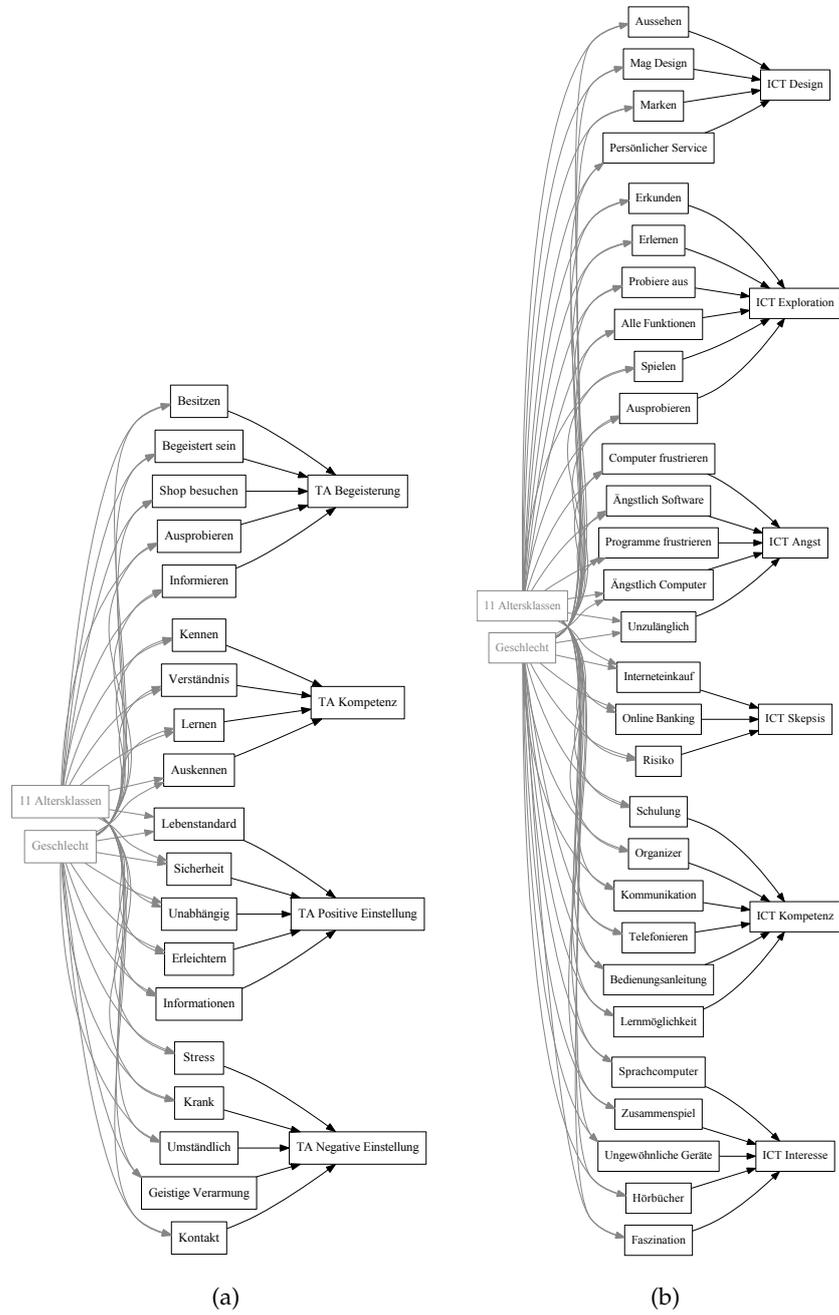


Abbildung 20: Bayes'sches Netz: Technikaffinität und Technikeinstellung (Teilnetze)

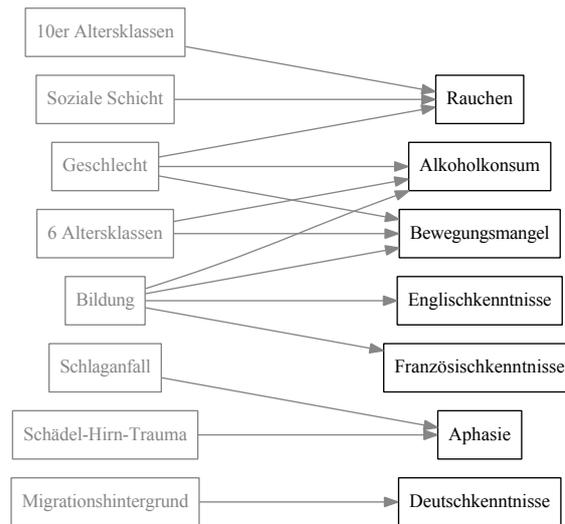


Abbildung 21: Bayes'sches Netz: Weitere Attribute (Teilnetz)

det werden, der die Technikaffinität anhand von 19 Items auf einer 5-stufigen Likert-Skala erfasst. Da die einzelnen Dimensionen nur etwas über 50 % der Varianz in den einzelnen Items erklären werden, werden zur Abbildung der Technikaffinität im Bayes'schen Netz alle 19 Items modelliert. Jeder Knoten hat 5 Ausprägungen, die die Antwortmöglichkeiten auf der entsprechenden Likert-Skala abbilden. In Abbildung 20a sind die Zusammenhänge zwischen den Knoten dargestellt.

Für die Persönlichkeitseigenschaft *Technikeinstellung* wird auf die Arbeit von Weiss et al. [247] zurückgegriffen. Mit der Technikeinstellung wird die Einstellung einer Person gegenüber Informations- und Kommunikationstechnologie als sechsdimensionales Konstrukt beschrieben, das über einen Fragebogen mit 29 Items erfasst werden kann. Wie schon bei der Technikaffinität, werden allen Items des Fragebogens modelliert und pro Dimension mittels einer beschränkender Variable verbunden (siehe Abbildung 20b).

Da weder für Technikaffinität noch für Technikeinstellung statistisch gesicherte Daten über die Verteilung in der Gesellschaft existieren, dienen als Datenquellen die über die Jahre im Institut erhobenen Fragebögen, welche unterschiedlichen Studien entstammen.

3.2.2.8 Weitere Attribute

In diesem Abschnitt über weitere Attribute werden Einschränkungen und Risiken beschrieben, die die Sprache (Aphasie und Sprachkenntnisse) und allgemeine Risikofaktoren (Alkoholkonsum, Rauchen und wenig Bewegung) betreffen. Der Zusammenhang zwischen den einzelnen Attributen ist in Abbildung 21 dargestellt.

Bei einer Aphasie handelt es sich um eine erworbene Sprachstörung, die ursächlich aus einer Erkrankung des zentralen Nervensystems folgt [126, S. 88]. Die häufigsten Ursachen für eine Aphasie sind der Schlaganfall oder das Schädel-Hirn-Trauma. Obwohl sich Aphasien in verschiedene Formen einteilen lassen, wird der Knoten binär modelliert, da epidemiologische Zahlen für die einzelnen Hauptarten (*Broca Aphasie*, *Wernicke Aphasie*, *Amnestische Aphasie* und *globale Aphasie*) fehlen.

Da die Verteilung der Sprachkenntnisse vom Bildungsniveau abhängt, wurden die Sprachknoten (*Englischkenntnisse*, *Französischkenntnisse*) mit dem Knoten *Bildung* verbunden. Bei den Sprachkenntnissen ist die Datenlage in Deutschland sehr beschränkt. Es existieren Statistiken über die Anzahl der Schüler, die eine bestimmte Fremdsprache z. B. Englisch in der Schule lernen [222], allerdings fehlen die Angaben über das erreichte Sprachniveau. Für Englisch und Französisch existiert eine Studie [109], in welcher die Befragten eine Selbsteinschätzung hinsichtlich ihrer Sprachkenntnisse angeben, allerdings sind die Angaben nur wenig detailliert, weshalb diese Knoten mit 3 Ausprägungen (*sehr gute*, *gute* und *geringe/keine Kenntnisse*) modelliert werden. Für Italienisch und Spanisch konnten keine Daten gefunden werden, weshalb die Knoten nicht modelliert wurden. Zur Modellierung der Deutschkenntnisse von Migranten wurde auf den Bericht von Haug [90] zurückgegriffen, wobei der Knoten binär (*vorhanden* oder *nicht vorhanden*) modelliert ist, nicht zwischen Kindern und Erwachsenen unterschieden wird und für Deutsche vereinfacht angenommen wurde, dass alle Deutsch sprechen und verstehen können.

Der Knoten *Alkoholkonsum* hat 5 Ausprägungen (*kein*, *risikoarmer*, *risikanter Alkoholkonsum*, *Alkoholmissbrauch* und *Alkoholabhängigkeit*). Die Verbindungen zwischen 6 *Altersklassen/Geschlecht* und *Alkoholkonsum* reflektieren den Umstand, dass insbesondere jüngere Männer mehr Alkohol konsumieren. Mit der Verbindung zwischen *Bildung* und *Alkoholkonsum* wird der Zusammenhang zwischen einem hohen Bildungsniveau und einem hohen Alkoholkonsum hergestellt.

Beim Knoten *Rauchen* wird zwischen *Nichtrauchern*, *Gelegenheitsrauchern* und *Starkrauchern* unterschieden. Die Unterscheidung zwischen Gelegenheitsraucher und Starkraucher geschieht anhand der Anzahl der gerauchten Zigaretten pro Tag, so ist nach der Definition der WHO eine Person als starker Raucher zu klassifizieren, wenn sie mehr als 20 Stück pro Tag raucht [144]. Bei beeinflussenden Faktoren zeigt sich, dass Männer stärker als Frauen, dass bei den Jüngeren mehr als bei den Älteren und dass bei den Personen aus den unteren Bildungsschichten mehr als bei Personen aus den anderen Bildungsniveaus geraucht wird.

Neben dem Alkoholkonsum und dem Rauchverhalten ist der Bewegungsmangel ein dritter Faktor, der zu einem höheren Krankheitsrisiko führt. Für die Modellierung wird eine binäre Ausprägung angenom-

men. Nach dem Studienbericht von Lange [149, S. 119 ff.] betätigen sich rund 40 % der Erwachsenen einmal pro Woche körperlich. Rund 20 % bewegen sich mindestens 5-mal pro Woche für mindestens 30 Minuten so, dass Sie ins Schwitzen kommen. Bei Befragten aus den unteren Bildungsschichten geben mehr Personen an, sich mehr häufiger körperlich zu bewegen, was an der Arbeitstätigkeit liegen kann. Ab 65 Jahren nimmt die Zahl derer, die sich aktiv körperlich bewegen, stark ab.

Für die Modellierung der Verteilungen in den Knoten werden verschiedene Bücher [108, 126] und Studien [141, 177] genutzt. Weiterhin werden Daten aus der Datenbank zur Gesundheitsberichtserstattung des Bundes [224] verwendet.

3.3 GENERIERUNG VON NUTZERMODELLEN

In diesem Abschnitt der Arbeit wird beschrieben, wie die Nutzermodelle mit Hilfe des Bayes'schen Netzes generiert werden können. Im ersten Unterabschnitt werden die Schritte beschrieben, die während des Ablaufs der Generierung durchlaufen werden. Hieran anschließend werden die Ergebnisse von Beispielabfragen beschrieben, um den Mehrwert der Lösung darzustellen und deren Validität zu verifizieren.

3.3.1 *Ablauf*

Wie bereits beschrieben, wird als Software für die Umsetzung die Implementierung von Druzdzel [56] verwendet. Hierdurch kann das Bayes'sche Netz auf zwei verschiedene Arten genutzt werden. Zum einen kann die Beschreibung des Netzes in die Software (GeNIe und SMILE) selbst geladen werden, um die Auswirkungen und Zusammenhänge der einzelnen Attribute zu sehen. Zum anderen können die Nutzermodelle über eine Schnittstelle in einer eigens programmierten Serverimplementierung erzeugt werden. Dieser Weg wird durch die MeMo-Werkbank verwendet und nachfolgenden beschrieben.

Um Nutzermodelle für die Werkbank erzeugen zu lassen, fragt die Werkbank über ein API den Server an. Beim Aufruf durch die Werkbank werden die Eigenschaften der Nutzermodellgruppe an den Server übermittelt. Die übermittelten Eigenschaften werden als Evidenz im Netz markiert. Durch die Einbringung der Evidenz, sind nun zwei weitere Operationen möglich.

Zum einen kann durch die Software die Wahrscheinlichkeit für das Vorliegen der Evidenz ($P(e)$) berechnet werden. Es kann somit Aussagen über die Abdeckung der möglichen Fälle und deren Relevanz getroffen werden. Zum anderen dient die Evidenz dazu, zu bestimmen, ob ein durch Sampling erzeugtes Nutzermodell verworfen werden muss oder nicht. Für das Sampling von Nutzermodellen werden die Knoten des Netzes basierend auf den Kanten in einer Liste

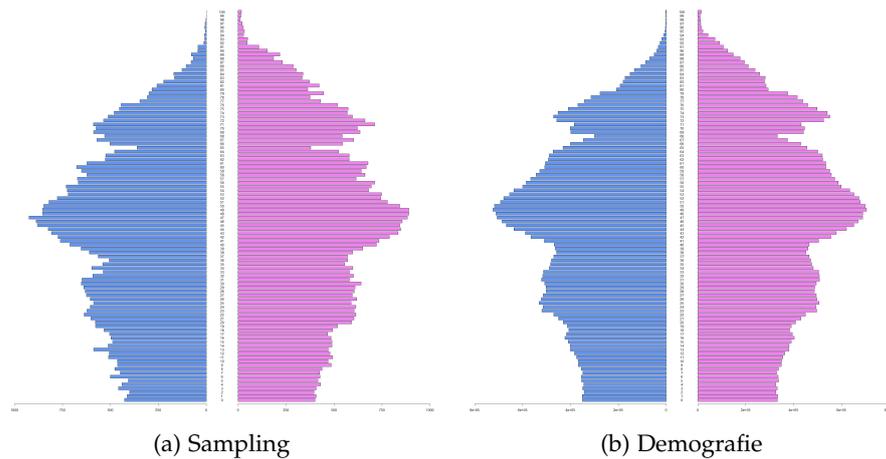


Abbildung 22: Alters- und Geschlechterverteilung

geordnet, so dass Elternknoten vor den Kinderknoten stehen. Durch einen Zufallsalgorithmus wird eine Ausprägung für jeden Knoten bestimmt. Stimmen die Ausprägungen der Knoten nicht mit der Evidenz überein, so wird das aktuelle Nutzermodell verworfen, im anderen Fall wird die generierte Ausprägung der Knoten (die Beschreibung des Nutzermodells) in einem Pool gespeichert. Ist die benötigte Anzahl von Nutzermodellbeschreibungen erzeugt, so werden diese Beschreibungen durch den Server in ein JSON-Format umgewandelt, welches an die MeMo-Werkbank übertragen werden kann.

In der MeMo-Werkbank werden die Nutzermodellbeschreibungen in konkrete Nutzermodelle umgewandelt und für die Simulation verwendet.

3.3.2 Beispiele und Validierung

In diesem Abschnitt werden verschiedene Beispiele dargestellt, die zeigen, wie das Bayes'sche Netz verwendet werden kann und welchen Mehrwert die Verwendung bietet. Für jedes Beispiel wird beschrieben, wie die Rahmenbedingungen sind, welches Ziel verfolgt wird und wie die Ergebnisse aussehen.

3.3.2.1 Beispiel 1

Im ersten Beispiel wird keine Evidenz (keine Einschränkung bzw. Limitierung) ins Netz eingebracht, bevor die Nutzermodelle hieraus generiert werden. Das Ziel des Beispiels ist es, Nutzermodelle zu generieren, die demografische Verteilungen in der Gesellschaft widerspiegeln.

In [Abbildung 22](#) sind die Alters- und Geschlechterverteilungen (die Alterspyramiden) für die durch Sampling aus dem Bayes'schen Netz gewonnenen Daten ([Abbildung 22a](#)) und die demografischen Daten

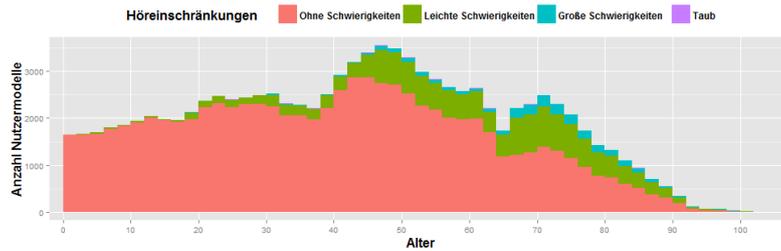


Abbildung 23: Ergebnis des Samplings für die Verteilung von Höreinschränkungen in verschiedenen Altersabschnitten

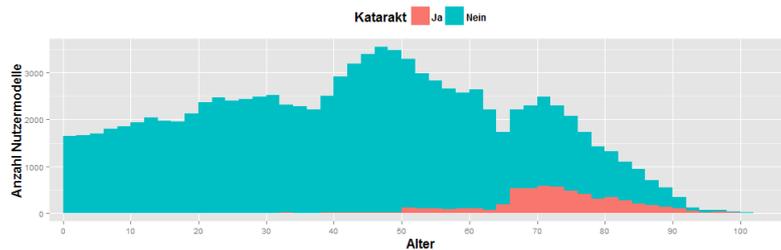


Abbildung 24: Ergebnis des Samplings für die Verteilung von Katarakten in verschiedenen Altersabschnitten

(Abbildung 22b) dargestellt. In der Abbildung sind die Anteile der Frauen an der Gesellschaft als lila Balken dargestellt, während die Anteile der Männer mit blauen Balken visualisiert werden.

Wie anhand der Abbildungen zu sehen ist, kann die Alters- und Geschlechterverteilung durch Sampling aus dem Netz wiedergegeben werden. Die Alterspyramide die auf den Daten des Bayes'schen Netzes beruht ist weniger glatt, da nur insgesamt 100.000 Nutzermodelle aus dem Netz generiert wurden. Die leichte Verschiebung in der Altersstruktur resultiert aus den unterschiedlichen Zeitpunkten, die den Datenquellen zugrunde liegen. Die Daten für das Bayes'sche Netz entstammen dem Zensus 2011. Die demografischen Daten spiegeln den Stand der Bevölkerung zum Jahresende 2013 wider, wobei es sich hierbei um eine auf den Zensusdaten beruhende Abschätzung handelt. Die Quelle der demografischen Daten ist die GENESIS-Datenbank des Statistischen Bundesamts [223].

In den Abbildungen 23 und 24 sind beispielhaft zwei Verteilungen für Einschränkungen dargestellt, wie sie durch die Generierung von Nutzermodellen mit Hilfe des Bayes'schen Netzen gefunden wurden. In Abbildung 23 ist zu erkennen, dass der Anteil der Personen mit Höreinschränkungen mit zunehmendem Alter steigt. Bei den Nutzermodellen unter 17 Jahren sind weniger als 2 % von einer Höreinschränkung betroffen, wohingegen bei den Nutzermodellen über 65 Jahren der Anteil ohne Einschränkungen nur knapp 55 % beträgt. Ein ähnliches Bild zeigt sich bei den Katarakten (Abbildung 24). Wie in

Abschnitt 3.1.2 gezeigt, sind von Katarakten meist ältere Menschen betroffen. Bei den durch das Sampling generierten Nutzermodellen zeigt sich ebenfalls dieser Umstand. Bei den Nutzermodellen unter 60 Jahren ist nur jedes hundertste Nutzermodell von dieser Einschränkung betroffen. Bei den Modellen die für Personen stehen, die 90 Jahre oder älter sind, hat rund jedes dritte Nutzermodell eine Einschränkung des Visus durch einen Katarakt.

Mit Hilfe dieses Beispiels konnte gezeigt werden, dass das Netz in der Lage ist, die modellierten Daten korrekt bzw. plausibel wiederzugeben. Werden in einer Simulation diese so erzeugten Nutzermodelle verwendet, so kann bei entsprechender Verarbeitung der Einschränkungen ein realistischeres Simulationsergebnis erzielt werden.

3.3.2.2 *Beispiel 2*

Das zweite Beispiel deckt den Fall ab, dass das erzeugte Nutzermodell keine Einschränkungen hat. Es wird demzufolge die Annahme getroffen, dass das Nutzermodell normal sehen (weder Kurz- noch Weitsichtig) und hören kann, keine Behinderung hat und auch keine die Interaktion beeinflussende Erkrankung (Tremor, Gicht, usw.) vorliegt; ein Fall, der in der vorliegenden MeMo-Werkbank häufig auftritt. Das Ziel dieses Beispiels ist es zu zeigen, dass die bisher in der MeMo-Werkbank verwendeten Nutzermodelle nicht den demografischen Verteilungen entsprechen bzw. dass die bisher verwendeten Nutzermodelle nur einen kleinen Teil der Bevölkerung repräsentieren.

Wird die Annahme, dass keine Einschränkung vorliegt, als Evidenz in das Netz eingepflegt, so zeigt sich, dass nur rund 25 % der Bevölkerung beschrieben werden und dass die beschriebenen Personen zum überwiegenden Teil (> 80 %) der Gruppe der Kinder und Jugendlichen (6 bis 14-jährigen Kinder & Jugendliche) entstammen. Wird zusätzlich zu den genannten Annahmen die Festlegung getroffen, dass es sich bei dem Nutzermodell um eine Person handelt, die über 65 Jahre alt ist, so werden nur 0,2 % der Bevölkerung simuliert bzw. nur 1 % der Senioren durch die Simulation repräsentiert.

Das Bayes'sche Netz kann somit genutzt werden, um die Validität bzw. die Repräsentativität der Annahme zu testen.

3.3.2.3 *Beispiel 3*

Im dritten Beispiel wird Evidenz ins Netz gegeben und es wird betrachtet, wie sich die Wahrscheinlichkeitsverteilungen durch das Hinzufügen der Information verändern. Für dieses Beispiel wird die Information ins Netz gegeben, dass die Personen, die durch die Nutzermodelle repräsentiert werden sollen, einen Hauptschulabschluss haben. Wie in Abschnitt 3.1.1.2 wurde gezeigt, haben Personen aus den unteren Bildungsschichten ein höheres Risiko für Höreinschränkungen. Wie in Abschnitt 3.2.2.4 dargestellt, existiert keine direkte Verbindung

Schulabschluss	Höreinschränkung			
	keine	leichte	schwere	taub
Nicht definiert	81 %	16 %	3 %	0 %
Hauptschulabschluss	67 %	26 %	6 %	0 %
Abitur	87 %	11 %	1 %	0 %

Tabelle 21: Wahrscheinlichkeitsverteilungen für eine Höreinschränkung in Abhängigkeit vom Schulabschluss

zwischen den Knoten Schulabschluss und Höreinschränkungen. Die Hypothese für dieses Beispiel lautet: Der Anteil der Personen/Nutzermodelle mit Einschränkungen im Hören ist bei Nutzermodellen mit niedrigen Schulabschlüssen höher, als bei Nutzermodellen mit höherer Bildung. Das Ziel dieses Beispiels ist es zu zeigen, dass mit dem Netz auch Zusammenhänge dargestellt werden können, die nicht explizit modelliert wurden.

In der Tabelle 21 sind die Verteilungen für das Vorliegen einer Höreinschränkung aufgeführt. Im ersten Fall (*Nicht definiert*) ist keine Evidenz ins Netz hinzugefügt worden, es ist demnach nicht definiert, welchen Schulabschluss ein Nutzermodell hat. Im zweiten Fall ist die Information ins Netz eingefügt worden, dass die von den Nutzermodellen repräsentierten Personen als höchsten Schulabschluss einen Hauptschulabschluss haben. Wie zu erkennen ist, steigt der Anteil der Personen, die eine Höreinschränkung haben stark an. Wird, wie im dritten Fall angenommen, die Information ins Netz gegeben, dass die Nutzermodelle ein Abitur als höchsten Abschluss haben, so steigt der Anteil der Personen ohne Höreinschränkungen auf 87 % an.

Bei den Nutzermodellen mit niedrigem Schulabschluss ist der Anteil der Modelle mit Höreinschränkungen 2,67 mal so hoch, wie der Anteil der Modelle mit Einschränkungen im Hören bei den Nutzermodellen mit Abitur. Dies entspricht in etwa dem von Cruickshanks et al. [43] gefundenen Faktor von 2,4 (siehe Abschnitt 3.1.1). Es kann somit festgehalten werden: Mit dem BN können Zusammenhänge generiert werden, die nicht implizit modelliert wurden.

3.4 FAZIT

Die in diesem Kapitel vorgestellte Lösung ist in der Lage Nutzermodelle zu erzeugen, die in ihrer Gesamtheit und mit ihren individuellen Eigenschaften, die demografische Verteilungen in der Gesellschaft widerspiegeln. Durch die Auswahl der verwendeten Knoten können Verteilungen von Einschränkungen in der Wahrnehmung, Kognition und Motorik beschrieben werden. Weiterhin können Zusammenhänge zwischen verschiedenen Einschränkungen bzw. Nutzermodellattribu-

ten korrekt durch das Netz abgebildet werden, was die Verwendung von unrealistischen Annahmen in der Simulation verhindert.

3.4.1 *Positive Aspekte*

Mit dem Bayes'schen Netz ist es möglich die Verteilung von Einschränkungen in der Bevölkerung demografisch valide abzubilden (siehe Abschnitt 3.3.2.1). Durch die Definition von Abhängigkeiten (Kanten) im Netz werden die Zusammenhänge zwischen Nutzerattributen und Einschränkungen adäquat wiedergegeben. Der große Vorteil der vorgestellten Lösung ist, dass der Verwender keinerlei Fachkenntnisse über Verteilungen von Einschränkungen in der Bevölkerung haben muss, um entsprechende Aussagen über Verteilungen treffen zu können.

Gleichfalls ist durch die Verwendung des Bayes'schen Netzes eine hohe Flexibilität gegeben. Statt auf vorgegebene Nutzermodelle (Nutzermodellpool) zurückzugreifen, können individuelle Modelle in beliebiger Anzahl durch das Netz generiert werden. Der Verwender kann seine Annahmen (Evidenz) in das Netz einfügen und erhält basierend auf diesen Annahmen zum einen demografisch korrekte Nutzermodelle und zum anderen Informationen über die Wahrscheinlichkeit der Evidenz (siehe Abschnitt 3.3.2.2). Durch die Angabe der Wahrscheinlichkeit der Evidenz, kann der Verwender seine Annahmen überprüfen und gegebenenfalls anpassen. Durch diese Möglichkeit können unrealistische Nutzerbeschreibungen gefunden werden.

Durch Inferenz ist es möglich Schlüsse aus vorliegenden Informationen zu ziehen und bedingte Wahrscheinlichkeitsverteilungen anderer Knoten zu berechnen. Mit Hilfe der Inferenz können Zusammenhänge gefunden werden, die nicht im Netz modelliert wurden (siehe Abschnitt 3.3.2.3)

3.4.2 *Negative Aspekte*

Die hier vorgestellte Lösung hat neben den genannten positiven Aspekten auch negative Aspekte bzw. Unzulänglichkeiten, die genannt werden sollen. Zum einen besteht das vorgestellte Netz bisher aus nur 111 Knoten, wobei 23 für die Darstellung der Technikaffinität und 35 zur Abbildung der Technikeinstellung (ICT) verwendet werden, zum anderen konnten nicht alle bisher in der MeMo-Werkbank enthaltenen Attribute (Spanisch und Italienisch) aufgrund von fehlenden Daten aufgenommen werden.

In die gleiche Richtung zielt die Limitierung der Kanten im Netz. Wie in Abschnitt 3.1 gezeigt, ist das Zusammenspiel zwischen Einschränkungen hochkomplex. Leider konnte diese Komplexität nicht immer im Netz abgebildet werden. Einerseits, weil die Quantität der Zusammenhänge noch nicht erforscht ist oder andererseits weil die Daten hierfür fehlten.

Der dritte negative Aspekt betrifft die im Netz enthaltenen Daten. Bisher enthält das Netz nur Daten für Deutschland. Um eine breite Anwendung zu fördern und den Wert der Lösung weiter zu steigern ist es wünschenswert, dass weitere Länder und Kontinente modelliert werden, um die Unterschiede in den Verteilungen in den Ländern abbilden zu können.

Das Ziel dieses Kapitels ist die Auswirkungen von unpassenden Device Mental Model (DMMs) auf die Interaktion zu simulieren. Zu diesem Zweck wird für eine Geräteklasse (Touchscreen) dargestellt, wie die Auswirkungen von unpassenden mentalen Modellen auf die Interaktion erfasst, in ein Modell überführt und simuliert werden können.

Zu Beginn des Kapitels wird der empirische Versuch beschrieben, welcher der Modellbildung zugrunde liegt. Die Beschreibung des Versuchs umfasst den Versuchsaufbau, die teilnehmenden Versuchspersonen, den Ablauf des Versuches, die Angaben zur Annotation der Videos und die Präsentation der Ergebnisse des Versuchs. Da die Ergebnisse die Modellbildung beeinflussen, wird detailliert beschrieben, wie die Versuchspersonen mit dem System interagierten, welche Probleme dabei auftraten, und was die Gründe für die Probleme waren. Nach der Darstellung der Ergebnisse werden die wichtigsten Ableitungen zur Modellbildung kurz zusammengefasst. Die Beschreibung der Erweiterung der Werkbank schließt sich der Versuchsbeschreibung an. In diesem Abschnitt des Kapitels wird dargestellt, welche Teile der Werkbank erweitert bzw. verändert werden mussten, um das im Versuch beobachtete Verhalten reproduzieren zu können. Den Abschluss des Kapitels bilden die detaillierte Beschreibung der Simulationsbedingungen und der Vergleich zwischen den empirischen Daten und den simulierten Ergebnissen der erweiterten MeMo-Werkbank.

4.1 EMPIRISCHER VERSUCH ZUR MODELLBILDUNG

Der Versuch wurde durchgeführt, um zu untersuchen, wie ältere Personen in intelligenten Wohnumgebungen (Smart Home) interagieren und welche Probleme während der Interaktion auftreten. Wie in Abschnitt 1.3.3 gezeigt, ist es für ältere Personen schwierig mit modernen Eingabegeräten ein System zu steuern, da es vorkommen kann, dass die mentalen Modelle der Person nicht zu den verwendeten Eingabegeräten passen. In modernen intelligenten Wohnumgebungen können meist verschiedene Eingabemodalitäten genutzt werden, um die angeschlossenen Geräte zu steuern. Um netzwerkfähige Geräte in einem Smart Home zu steuern wird ein System zur Verarbeitung von Nutzereingaben benötigt. Im vorliegenden Versuch wurde das INSPIRE-System zur Steuerung der intelligenten Umgebung verwendet.



Abbildung 25: Versuchsumgebung.

Das System INSPIRE (INfotainment management with SPEech Interaction via REmote microphones and telephone interfaces) ermöglicht die Steuerung von netzwerkfähigen Haushaltsgeräten via Sprache [165]. Für den Versuch wurde das System erweitert, so dass eine Steuerung via Sprache, Smartphone und 3D-Gestensteuerung möglich war. Um Fehler bei der Sprach- und Gestenerkennung zu vermeiden, wurden der Spracherkennung und das System zur Gestenerkennung durch den Versuchsleiter (Wizard) simuliert.

4.1.1 Ziele

Das ursprüngliche Ziel des Versuchs war es herauszufinden, welche Eingabemodalität (*Sprache*, *Touch* oder *Gestensteuerung*) für jüngere und ältere Nutzer geeignet ist und welche Eingabemodalität von den Versuchspersonen präferiert wird.

Während des Versuchs zeigte sich, dass speziell ältere Probanden bei der Fernbedienung von Geräten mittels iPhone, Interaktionsmuster anwendeten, welche auf ein falsches bzw. nicht passendes mentales Modell hinsichtlich der Benutzung eines Smartphones schließen ließen. Da diese falschen mentalen Modelle das DMM betreffen, wurde beschlossen, die Ergebnisse des Versuches für den Entwurf der Modelle für die MeMo-Werkbank zu verwenden. Da nur die Touch-Interaktionen für die Simulation von DMM-Fehlern relevant sind, werden nachfolgend nur die Aspekte des Versuches erläutert, die für die Reproduzierbarkeit notwendig sind und die im weiteren Verlauf des Kapitels Verwendung finden.

4.1.2 Aufbau und Geräte

Der Versuch wurde in einem, zur intelligenten Wohnumgebung umgebauten, Labor durchgeführt (Abbildung 25). In der Versuchsumgebung konnten während des Versuchs die Rollläden, die Decken-

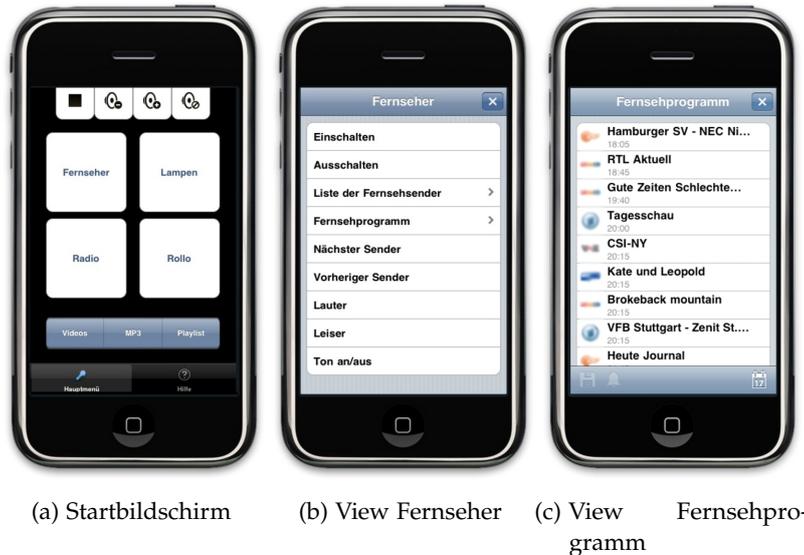


Abbildung 26: iPhone App

und Wandlampen, der Fernseher und das Radio durch das INSPIRE-System gesteuert werden. Während der Versuchsdurchführung saßen die Versuchspersonen auf dem Sofa, welches auf der rechten Seite in Abbildung 25 dargestellt ist. Vor den Versuchspersonen, auf dem Tisch, stand ein Laptop, auf welchem die aktuelle Aufgabenstellung angezeigt wurde. Die Versuchspersonen konnten von ihrem Sitzplatz direkt auf den Fernseher schauen. Der Abstand zwischen Sofa und Fernseher beträgt etwa 3 Meter. Unter dem Fernseher im Fernseh-Rack befand sich das Radio. Die Rollläden befanden sich rechts, die Wandlampen links vom Sofa.

Unter dem Fernseher war eine Kamera positioniert, welche die Interaktion der Versuchsperson mit dem System aufgezeichnet hat. Vor der Tür (abgebildet auf der linken Seite in Abbildung 25) war eine weitere Kamera aufgestellt, die es dem Wizard ermöglichte die Gesten der Versuchsperson zu erkennen. Um das System mit der Sprache zu bedienen, wurde die Versuchsperson mit einem Ansteckmikrofone ausgestattet. Nach Abschluss des Versuchsdurchgangs wurden die Versuchspersonen über den Zweck der zweiten Kamera informiert, weiterhin wurde ihnen mitgeteilt, dass die Gesten- und Spracherkennung durch einen Mitarbeiter ausgeführt wurden.

Um das System per Touch zu bedienen wurde ein iPhone 3GS eingesetzt. Auf dem iPhone wurde eine eigens entwickelte App installiert. Der Startbildschirm der App (siehe Abbildung 26a) bietet Zugriff auf 4 Untermenüs zur Bedienung des Fernsehers, der Lampen, des Radios und der Rollos. Wird einer dieser Button aktiviert zeigt das Smartphone eine neue Ansicht (*View*) an, welche gerätespezifische Aktionen ermöglicht. Die 4 Buttons am oberen Rand des Bildschirms ermöglichen die Wiedergabe zu stoppen, die Lautstärke zu verringern

bzw. zu erhöhen oder den Ton stumm zu schalten. Wird einer der Button gedrückt, erscheint ein neuer View mit der Auswahl der möglichen Geräte, auf die die gewählte Aktion angewendet werden kann. Unter den 4 großen Button befinden sich 3 kleinere Button die einen Schnellzugriff auf weitere Funktionen des Systems ermöglichen. So kann durch diese Button auf die Liste der aufgezeichneten Sendungen (*Videos*), die Musik (*MP3*) und die Wiedergabeliste (*Playlist*) zugegriffen werden. Am unteren Rand des Bildschirms befinden sich zwei Buttons, die einen Wechsel zwischen der Bedienoberfläche und der eingebauten Hilfefunktion ermöglichen.

Die Abbildung 26b zeigt den View *Fernseher*, welcher die Optionen anzeigt, die dazu genutzt werden können, den Fernseher zu bedienen. Den View aus Abbildung 26b erreicht man, indem man auf dem Startbildschirm den *Fernseher* Button berührt. Der Button in der rechten oberen Ecke in Abbildung 26b schließt den View und führt zur Anzeige des Startbildschirms. Die Button mit einem Größer-als-Zeichen (>) öffnen einen neue View mit weiteren Optionen oder Informationen. In Abbildung 26b führt das Berühren des Buttons *Fernsehprogramm* > dazu, dass der View in Abbildung 26c geöffnet wird. Die im Versuch eingesetzte App hatte leider einen Usability-Fehler, so führte das Ein- bzw. Ausschalten des Fernsehers oder ein Kanalwechseln dazu, dass der aktuelle View *Fernseher* geschlossen und stattdessen der Startbildschirm (Abbildung 26a) angezeigt wurde.

Die Abbildung 26c zeigt den View, welches das Fernsehprogramm anzeigt. In diesem View ist es möglich einzelne Programme aufzeichnen zu lassen oder sich zu Beginn einer Sendung an diese erinnern zu lassen. Der Button rechts oben mit dem X schließt den aktuellen View und führt dazu, dass der View aus Abbildung 26b angezeigt wird. Die Bedienung des Fernsehprogramms war nicht Teil der Aufgabenstellung, deshalb wird im Folgenden nicht näher darauf eingegangen.

Während des Versuches und in der Gewöhnungsphase wurden die Eingaben per *Arm-Geste* bzw. *Sprache* in das System, als auch die Interaktionen auf dem Smartphone, in Logfiles aufgezeichnet. Die auf dem iPhone installierte App konnte nur Eingaben aufzeichnen, die eine Aktion auslösten, so dass nicht erkannte Interaktionen bzw. das Verfehlen eines Buttons nicht aufgezeichnet wurden. Um dieses Problem auszugleichen wurden die aufgezeichneten Videos später annotiert (siehe Abschnitt 4.1.6).

4.1.3 Versuchspersonen

Am Versuch nahmen 16 jüngere und 17 ältere Versuchspersonen teil. Die jüngeren Versuchspersonen (20 - 29 Jahre; Median = 24 Jahre; Mittelwert = 24,38; SD = 2,68; 8 Frauen) wurden auf dem Campus der TU Berlin rekrutiert. Von den jüngeren Versuchspersonen nutzen 5 Personen täglich ein Gerät, welches einen Touchscreen als Eingabe-

modalität bietet, 8 Personen 1 bis 2 Mal pro Woche und 3 Personen nutzen Geräte mit einem Touchscreen seltener als 1 Mal pro Woche.

In der Gruppe der älteren Versuchspersonen (52-67 Jahre; Median=58 Jahre; Mittelwert=59; SD=4,34; 9 Frauen) nutze eine Person täglich Geräte mit einem Touchscreen, 2 Personen nutzen ein Gerät mit Touchscreen 3 bis 6 Mal pro Woche, 8 Versuchspersonen 1 bis 2 Mal pro Woche und 6 Personen solch ein Geräte seltener als 1 Mal pro Woche. Die älteren Versuchspersonen wurden durch Aushänge in Supermärkten geworben.

Alle Versuchspersonen erhielten für ihre Teilnahme am Versuch eine Bargeldentschädigung in Höhe von 15,- Euro. Keine Versuchsperson kannte das eingesetzte System (INSPIRE) aus vorherigen Versuchen.

Für die nachfolgende Auswertung wurden die Daten von 5 älteren Versuchspersonen nicht betrachtet, weil die Logfiles fehlerbehaftet sind (2 Männer, 1 Frau), das eingesetzte System während des Versuches abstürzte (1 Mann) oder die Versuchsperson (1 Mann) nicht in der Lage war das eingesetzte iPhone zu bedienen. Für die Auswertung können somit Daten von 16 jüngeren und 12 älteren Versuchspersonen (54-67 Jahre, Median=57,5; Mittelwert=58,8; SD=3,97; 7 Frauen) genutzt werden.

4.1.4 Ablaufplan

Jeder Versuchsdurchgang des Experiments bestand aus 3 unterschiedlichen Blöcken. Nach der Begrüßung und Einweisung der Versuchsperson, wurden die Teilnehmer im ersten Block mit den möglichen Systemausgaben konfrontiert. Die Aufgabe für die Versuchspersonen bestand darin, die unterschiedlichen Systemausgaben (Aufnahmen der Sprachausgabe, Bilder der grafischen Oberflächen) zu bewerten. Für die Bewertung wurde der AttrakDiff mini [88] genutzt, der die Bewertung hinsichtlich der hedonischen Qualitäten (*Identität* und *Stimulation*), der *pragmatischen Qualität* und der *Attraktivität* auf einer 7-stufigen Skala ermöglicht.

Der zweite Block wurde genutzt, um die Eignung der einzelnen Modalitäten (*Sprache*, *Arm-Geste* und *Touch*) für die beiden Nutzergruppen zu testen. Alle Versuchspersonen mussten alle Eingabemodalitäten nacheinander testen und nach der Durchführung des Tests die Eignung der Modalität bewerten. Für die Bewertung wurde wiederum der AttrakDiff mini eingesetzt. Um Reihenfolgeeffekte zu vermeiden, wurde die Reihenfolge der Modalitäten zwischen den Versuchspersonen randomisiert. Jeder Test einer Modalität begann mit einer 5 minütigen Explorationsphase, welche von der Versuchsperson genutzt werden konnte, um sich an die Bedienung des Systems mit der festgelegte Modalität zu gewöhnen. Nach der Eingewöhnungsphase mussten die Versuchspersonen verschiedene Geräte (Rollo, Lampen, Radio

und Fernseher) bedienen und mit diesen Geräten definierte Aufgaben lösen.

Der dritte Block wurde genutzt, um die Präferenz der Versuchsperson für eine oder mehrere Modalitäten zu erfassen. In diesem Block war es den Teilnehmern freigestellt welche Modalität sie nutzen und welche Geräte sie mit der gewählten Modalität bedienen wollen.

Jeder Versuchsdurchgang endete damit, dass von der Versuchsperson zwei Fragebögen ausgefüllt wurde, die dazu dienten die Technikaffinität (TA; [127]) und Technikeinstellung (attitude towards information and communication technology; ICT; [247]) zu erfassen. Zusätzlich wurde, wie von Dickinson empfohlen [52], einen Zahlenspanntest [245] durchgeführt, um kognitive Einschränkungen im Kurzzeitgedächtnis zu erfassen.

4.1.5 Aufgabenstellungen

In diesem Abschnitt werden die Aufgabenstellungen erläutert, welche von den Versuchspersonen im unimodalen Block des Versuches (zweiter Versuchsblock) gelöst werden mussten. Da im weiteren Verlauf des Kapitels nur die Smartphone-Interaktionen betrachtet werden, wird die Beschreibung der Aufgaben anhand der Modalität *Touch* erklärt. Die Aufgabenstellungen sind jedoch auf die Arm-Gesten- und Sprachsteuerung zu übertragen. Die Reihenfolge der Aufgaben wurde nicht variiert, so dass die Versuchspersonen für alle Modalitäten immer erst das Rollo, anschließend die Lampen, hiernach das Radio und am Ende den Fernseher bedienen mussten.

Die erste Aufgabe im unimodalen Block des Versuches bestand darin die Rollläden zu steuern. Jede Versuchsperson musste die Rollläden herunterfahren, anhalten, herauffahren und stoppen. Mit dem Smartphone waren mindestens 5 Interaktionen notwendig. Erst musste die Versuchsperson den Button *Rollo* berühren (siehe Abbildung 26a), danach mussten nacheinander die Button *Rollo herunterfahren*, *Rollo anhalten*, *Rollo herauffahren* und *Rollo anhalten* bedient werden.

Die zweite Aufgabe bezog sich auf die Steuerung der Wand- und Deckenlampen. Mit dem Smartphone waren mindestens 5 Schritte notwendig. Erst muss die Versuchsperson den entsprechenden View öffnen, indem sie auf dem Startbildschirm den Button *Lampen* berührt. Danach muss die Versuchsperson das Deckenlicht einschalten, es dunkler machen, die Wandlampe heller machen und schlussendlich alle Lampen ausschalten.

Die Bedienung des Radios benötigte auf dem Smartphone mindestens 7 Schritte. Zur Lösung der Aufgabe musste die Versuchsperson den Button *Radio* auf dem Startbildschirm berühren. Auf dem neuen View musste die Versuchsperson das Radio einschalten, leiser machen, zum nächsten Sender wechseln, lauter machen, zum vorherigen Sender wechseln und schlussendlich das Radio ausschalten.

Zur Bedienung des Fernsehers musste die Versuchsperson als erstes den entsprechenden View aufrufen, indem sie auf dem Startbildschirm den passenden Button (*Fernseher*) berührt. Im View *Fernseher* mussten sie den Button *einschalten* bedienen, um den Fernseher einzuschalten. Aufgrund eines Usability-Fehlers wechselte nun das iPhone selbstständig in den initialen Startbildschirm, so dass zum Umschalten auf den nächsten Kanal wieder der View *Fernseher* aufgerufen werden musste. Im View *Fernseher* musste die Versuchsperson auf den Button *Nächster Sender* tippen, um einen Kanal weiter zu schalten. Wurde die Aktion vom System ausgeführt, so kam es auch hier zum schon bekannten Fehler, dass sich der View schloss und die Versuchsperson den Startbildschirm angezeigt bekam. Zur Reduzierung der Lautstärke konnte die Versuchsperson nun aus zwei unterschiedlichen Lösungspfaden wählen. Die erste Lösung verlangte, dass die Versuchsperson den View *Fernseher* öffnet und dort auf den Button *leiser* tippt. Anders als bei den vorherigen Aktionen wurde der View *Fernseher* nun allerdings nicht geschlossen. Die zweite Lösung verwendet die Buttons am oberen Rand des Startbildschirms. Hier musste die Versuchsperson erst auf den Button *leiser* tippen. Vom System wurde daraufhin ein neuer View erzeugt, welcher abfragt welches Geräte leiser gemacht werden soll. Nach der Auswahl des Gerätes wird wieder der Startbildschirm angezeigt. Zum Wechseln auf den vorherigen Kanal kommt es nun darauf an, welcher View angezeigt wird. Hat sich die Versuchsperson für den Lösungsweg 1 entschieden, so kann sie direkt auf den Button *Vorheriger Sender* tippen, um den Kanal zu wechseln. Hat die Versuchsperson den zweiten Lösungsweg gewählt, so musste sie erst den View *Fernseher* aufrufen und dort den entsprechenden Button auswählen. In beiden Fällen führt die Ausführung der Aktion dazu, dass sich der View *Fernseher* geschlossen und der Startbildschirm angezeigt wird. Die Lösungspfade für die Erhöhung der Lautstärke sind äquivalent zu den Lösungspfaden zur Reduzierung der Lautstärke, und für das Ausschalten des Fernsehers kommt es wiederum darauf an, welchen Lösungsweg die Versuchsperson zur Erhöhung der Lautstärke gewählt hat. Für die Bedienung des Fernsehers waren, bedingt durch den Usability-Fehler, mindestens 10 Schritte nötig.

4.1.6 Videoannotation der Smartphone-Interaktion

Da die benutzte iPhone-App keine Interaktionen aufzeichnen konnte, die nicht direkt zu einer Aktion im System führten, musste für jede Interaktion mit dem Smartphone annotiert werden, ob die Interaktion verarbeitet wurde oder nicht. Da nicht jeder Interaktionstyp gleich oft zu Erkennungsfehlern führte, wurde zusätzlich die *Intendierte Interaktion* der Versuchsperson annotiert.

Die Tabelle 22 zeigt das für die Annotation verwendete Schema. Die *intendierten Interaktionen* beschreiben die beabsichtigten Interaktionen

Attribut	Klassen
Intendierte Interaktion	Short tap, Short press, Long tap, Long press
Verarbeitete Interaktion	Short tap, Unbekannt
Interaktionsfehler	Mistake, Slip, Lapse, kein Fehler

Tabelle 22: Annotationsschema

der Versuchspersonen und unterscheiden sich anhand der Dimensionen *Zeit* und *aufgewendeter Druck*. Die Dimension *Zeit* beschreibt die Dauer, die der Finger der Versuchsperson den Bildschirm berührt. In der Literatur existiert keine Definition, wann eine Interaktion auf einem Smartphone als kurz (*short*) oder lang (*long*) definiert werden kann. Die Entwickler des Android-Betriebssystems definieren die Dauer, die nötig ist, um einen *Short tap* (kurze Berührung) als *Long tap* (lange Berührung) zu interpretieren, mit 500 ms [6]. Für die Annotation wurde angenommen, dass diese Zeitangabe auf Apple Smartphones übertragen werden kann, allerdings steht die Dauer der Interaktion nicht zur Verfügung, und so musste aus den Videos abgeleitet werden, ob es sich um eine kurze oder lange Interaktion handelt. Mit dem verwendeten Smartphone und der installierten App war es weiterhin nicht möglich den von der Versuchsperson aufgewendeten Druck zu messen, weshalb auch diese Dimension annotiert werden musste.

Hinsichtlich der verarbeiteten Interaktion ist eine halb-automatische Annotation möglich. Wurde vom iPhone eine Aktion auf einem Button erkannt, so wurde die Aktion als *Short tap* annotiert. Die Annahme, dass die Aktion als *Short tap* erkannt wurde, basiert auf dem Wissen, dass nur diese Eingabe als gültige Eingabe für die dargestellten Button definiert war. Wurde vom iPhone keine Aktion erkannt, im Video war jedoch eine Aktion der Versuchsperson zu erkennen, so wurde die verarbeitete Interaktion mit *Unbekannt* annotiert. Die allgemeine Klasse *Unbekannt* ist nötig, da aus den Videos nicht bestimmt werden kann, ob das iPhone eine andere Aktion (z. B. *Long tap*, *Scrolling* oder *keine Aktion*) verarbeitet hat.

Die Annotation des Typs des Interaktionsfehlers geschah nach der Annotation der intendierten Interaktion automatisch. Die Klassen für dieses Attribut leiten sich aus der Definition der Fehlertypen im Bereich Mensch-Maschine-Interaktion von Reason [186] ab. Nach Reason [186, S. 12 ff.] gibt es drei Typen von Interaktionsfehlern, namentlich *Mistakes*, *Lapses* und *Slips*.

In die erste Klasse (*Mistakes*) gehören Fehler, denen eine falsche Intention, ein fehlerhafter Plan zugrunde liegt. Ein Mistake betrifft das mentale Model. Bei einem Mistake wird einer Aktion ein erwartetes Ergebnis zugeordnet, welches nicht mit dem tatsächlich möglichen

Ergebnis übereinstimmt. Ein Beispiel für einen Mistake ist der Versuch einen Button auf einem kapazitiven Touchscreen mit einem Fingernagel zu bedienen. Da ein kapazitiver Touchscreen die Veränderung des elektrischen Leitwertes misst, der Fingernagel diesen Leitwert allerdings nicht verändert, kann der kapazitive Touchscreen keine Änderung messen, weshalb keine Aktion ausgelöst wird. Im Rahmen der Annotation wurden intendierte Interaktionen die kein *Short tap* waren, als *Mistake* klassifiziert.

Die zweite Klasse (*Lapse*) umfasst Fehler, die mit dem Erinnerungsvermögen der Person zusammenhängen. Der häufigste Fehler in dieser Klasse ist das Auslassen bzw. das Vergessen eines benötigten Schrittes, z. B. das Vergessen einer Teilaufgabe. Da Lapses in vorliegenden Fall keinen Einfluss auf das einzelne Interaktionsereignis haben, werden diese Fehler nachfolgend nicht weiter betrachtet. Im Gegensatz zu Mistakes und Lapses betreffen *Slips* die Ausführung der intendierten Aktion. Ein Beispiel für einen Slip ist das Verfehlen eines Buttons. Ein Slip ist immer unbeabsichtigt.

Für alle 28 Versuchspersonen wurden die Interaktionen mit dem Smartphone annotiert. Insgesamt wurden 1186 Interaktionen betrachtet. 120 Interaktionen (10,12 %) konnten nicht bewertet werden, da die Interaktionen auf den Videos nicht sichtbar waren (z. B. durch die Versuchsperson oder einen Gegenstand verdeckt). Von den auswertbaren Interaktionen (1066 Datenpunkte) wurden 232 intendierte Interaktionen von einem zweiten Experten ein weiteres Mal annotiert. Die Annotation durch einen zweiten Experten hatte das Ziel die Objektivität und Reliabilität der Annotation und die Eignung des Annotationsschemas zu testen.

Um den Grad der Übereinstimmung für die Annotation der intendierten Interaktionen zu bestimmen wurde das Cohens Kappa berechnet. Cohens Kappa ist eine statistische Methode, um die Übereinstimmung zwischen 2 Beurteilern auf einem Index abzubilden [37]. Cohens Kappa kann maximal einen Wert von 1 annehmen, was einer perfekten Übereinstimmung entspricht.

Für die Annotation der intendierten Interaktion konnte ein Kappa von 0,494 berechnet werden. Für die einzelnen Dimensionen *Zeit* und *aufgewendeter Druck*, sind die Kappa-Werte höher. So konnte für die Dimension *Zeit* ein Kappa-Wert von 0,665 und für die Dimension *ausgeübter Druck* ein Kappa Wert von 0,506 erzielt werden. Da alle Kappa-Werte größer als 0,4 sind kann nach Landis und Koch [145] von einer moderaten (Kappa > 0.4), für die Dimension *Zeit* sogar von einer guten Übereinstimmung (Kappa > 0.6) ausgegangen werden. Aufgrund der Kappa-Werte ist festzustellen, dass das Schema in Tabelle 22 zur Annotation der Attribute prinzipiell geeignet ist.

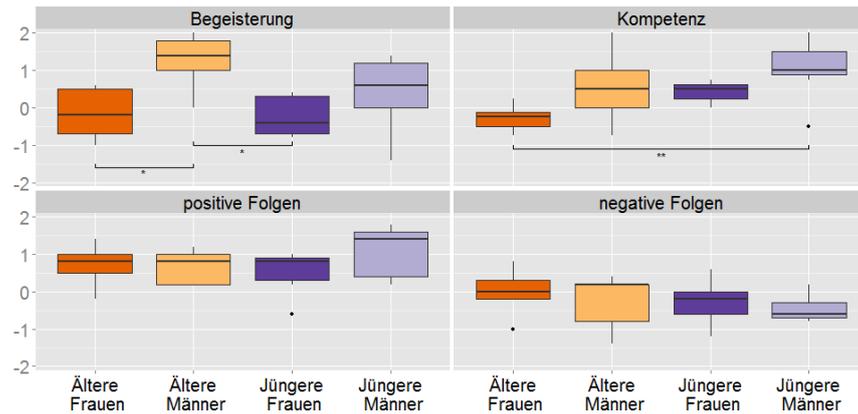


Abbildung 27: Verteilung in den Dimensionen der Technikaffinität für die 4 Versuchspersonengruppen

4.1.7 Ergebnisse

In diesem Abschnitt der Arbeit werden die Ergebnisse aus dem empirischen Versuch dargestellt. Im ersten Unterabschnitt werden die Ergebnisse der Technikaffinität (TA) und Technikeinstellung (ICT) für die Gruppen der jüngeren Frauen, älteren Frauen, jüngeren Männern und älteren Männern dargestellt. Die Fokussierung auf die genannten Eigenschaften erfolgt, da nur diese später in den allgemeinen linearen Modellen verwendet werden. Im Anschluss an die Nutzerattribute werden die Kennzahlen zu Interaktionshäufigkeiten und Interaktionsformen dargestellt. Es folgt der Unterabschnitt 4.1.7.3 welcher darstellt, wie häufig die Versuchspersonen fehlerhaft mit dem System interagiert haben. Denn Abschluss bildet die Darstellung der unnötigen Interaktionen.

4.1.7.1 Nutzerattribute

Die Abbildung 27 zeigt die Verteilungen für die einzelnen Dimensionen der Technikaffinität, wenn die Versuchspersonen nach Alter und Geschlecht in vier Versuchspersonengruppen eingeordnet werden.

Für die Dimension *Begeisterung für Technik* zeigte die einfaktorielle Varianzanalyse (ANOVA) signifikante Unterschiede ($F_{df(3,22)} = 4,391$; $p = 0,015$) zwischen den einzelnen Versuchspersonengruppen. Als Post-hoc-Test wurde der paarweise t-Test mit sequenzieller Bonferroni-Korrektur nach Holm genutzt. Die Korrektur des α -Wertes ist notwendig, um eine Alphafehlerkumulierung zu vermeiden. Die eingesetzte Methode ist als konservative einzustufen [102]. Das Ergebnis des Tests zeigt signifikante Unterschiede zwischen älteren Männern und älteren Frauen ($p = 0,028$) bzw. zwischen älteren Männern und jüngeren Frauen ($p = 0,022$). Ein Test der Effektstärke mittels Hedges' g (siehe Hedges [94]) zeigt einen großen Unterschied ($|g| = 1,755$) zwischen

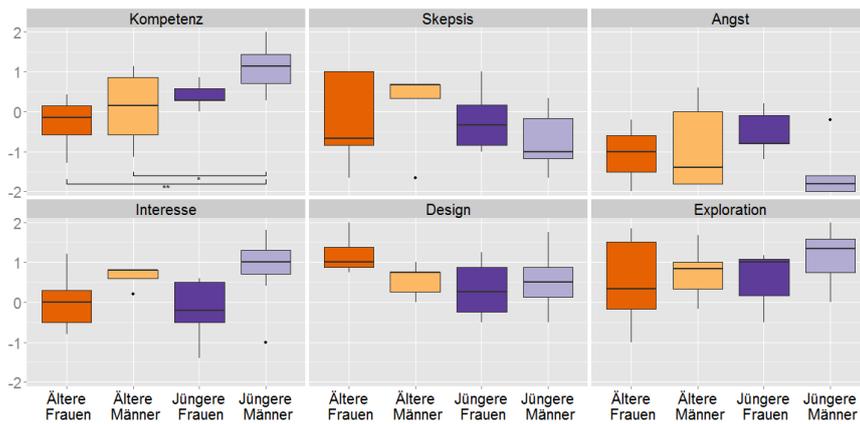


Abbildung 28: Verteilung in den Dimensionen der Technikeinstellung für die 4 Versuchspersonengruppen

älteren Frauen und älteren Männern und ebenfalls einen großen Unterschied ($|g| = 2,066$) zwischen jüngeren Frauen und älteren Männern. Man kann den am Versuch beteiligten älteren Männern somit eine hohe Begeisterung für Technik unterstellen. Da die älteren Männer sonst in keiner Dimension signifikante Abweichungen zeigen, wird davon ausgegangen, dass der Unterschied in der Begeisterung für Technik keinen Einfluss das Verhalten der Versuchsperson während der Interaktion mit dem System hat.

Für die Dimension *Kompetenz* zeigte die parametrische ANOVA ebenfalls signifikante Unterschiede ($F_{df(3,22)} = 4,807$; $p = 0,01$) zwischen den Versuchspersonengruppen. Der Post-hoc-T-Test (mit sequenzieller Bonferroni-Korrektur nach Holm) zeigte signifikante Unterschiede ($p = 0,006$) zwischen den älteren weiblichen Versuchspersonen und den jüngeren männlichen Versuchspersonen. Der Unterschied zwischen den jüngeren Männern und den älteren Frauen ist groß ($|g| = 1,997$). Der Unterschied kann darauf zurückgeführt werden, dass ältere Personen und Frauen sich weniger kompetent im Umgang mit Technik einschätzen (vgl. Karrer et al. [127]).

In der Dimension, die die positiven Folgen der Technik repräsentiert, zeigten sich, wie auch in der Dimension der negativen Technikfolgenabschätzung, keine signifikanten Unterschiede zwischen den Versuchspersonengruppen.

In Abbildung 28 ist die Verteilung der Versuchspersonen in den Dimensionen der Technikeinstellung (ICT) dargestellt. Für den ICT-Fragebogen konnten nur in der Dimension *Kompetenz* signifikante Unterschiede zwischen den Versuchspersonengruppen gefunden werden. Wie in Abbildung 28 markiert, wurden mit der parametrischen ANOVA hoch signifikante Unterschiede zwischen den Versuchspersonengruppen für die Dimension *Kompetenz* gefunden. Der eingesetzte Post-hoc-T-Test mit Bonferroni-Holm-Korrektur signalisiert, wie schon für den Technikaffinitätsfragebogen, hoch signifikante Unterschiede

Aufgabe	Ältere männliche Versuchspersonen	Jüngere männliche Versuchspersonen	Ältere weibliche Versuchspersonen	Jüngere weibliche Versuchspersonen
Rollo	M = 8,4 SD = 4,51	M = 6,57 SD = 0,79	M = 9 SD = 3,58	M = 6,57 SD = 1,62
Lampen	M = 6,4 SD = 2,7	M = 9,29 SD = 2,29	M = 7,17 SD = 3,25	M = 9,86 SD = 4,53
Radio	M = 12,2 SD = 3,19	M = 11,43 SD = 2,44	M = 11 SD = 3,37	M = 12,71 SD = 3,3
Fernseher	M = 20,4*** SD = 2,51	M = 11 SD = 1,15	M = 13,14 SD = 2,91	M = 13 SD = 3,27

Tabelle 23: Anzahl der benötigten Schritte

($p < 0,001$) zwischen den älteren weiblichen Versuchspersonen und den jüngeren männlichen Versuchspersonen. Zusätzlich zum erwarteten Unterschied zeigt sich hier noch ein signifikanter Unterschied zwischen den älteren männlichen Versuchspersonen und den jüngeren männlichen Versuchspersonen ($p = 0,05$). Bei der Bestimmung der Größe des Effektes zeigen sich sowohl für den Unterschied zwischen jüngeren Männern und älteren Frauen ($|g| = 2,171$) und jüngeren und älteren Männern ($|g| = 1,244$) große Effektstärken. Knapp nicht signifikante Unterschiede zwischen den Versuchspersonengruppen existieren in den Dimensionen Angst ($F_{df(3,22)} = 2,649$; $p = 0,074$), Design ($F_{df(3,22)} = 2,604$; $p = 0,078$) und Interesse an ICT ($F_{df(3,22)} = 2,897$; $p = 0,058$). Für die Dimensionen Skepsis und Exploration können keine signifikanten Unterschiede zwischen den Gruppen gefunden werden.

4.1.7.2 Interaktionen

Die im Versuch gestellten Aufgaben konnten, bei optimalem Interaktionsverlauf, mit 27 Interaktionsschritten gelöst werden. In Tabelle 23 sind die für die einzelnen Aufgaben, getrennt nach Versuchspersonengruppen, tatsächlich benötigten Schritte aufgeführt.

Der Vergleich der Mittelwerte über alle Versuchspersonengruppen für alle Teilaufgaben zeigte nur für den TV-Task signifikante Unterschiede zwischen den Gruppen ($F_{df(3,22)} = 13,86$; $p < 0,001$). Der als Post-hoc-Test durchgeführte paarweise T-Test mit Bonferroni-Holm-Korrektur führte zu dem Ergebnis, dass die Versuchspersonengruppe der älteren Männer im TV Task signifikant mehr Schritte benötigte, als dies für die anderen Gruppen der Fall war.

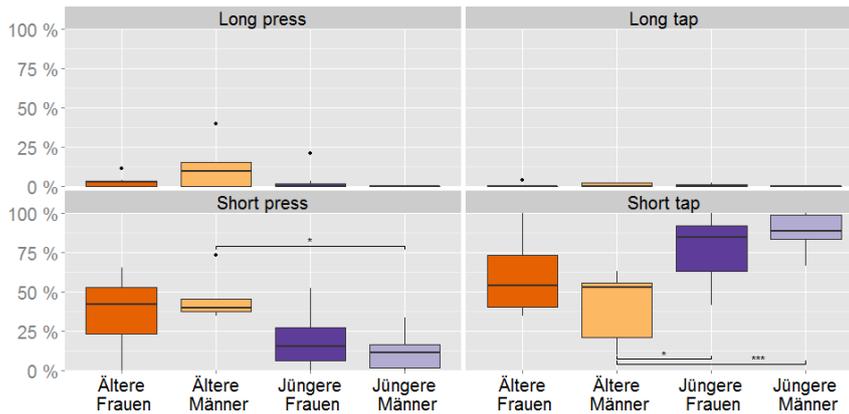


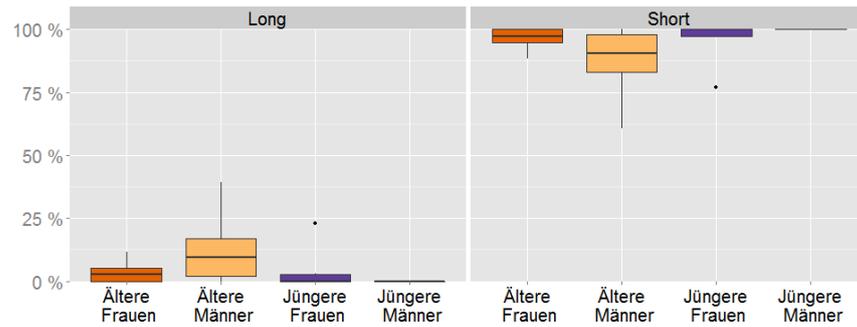
Abbildung 29: Häufigkeiten der intendierten Interaktion in den Versuchspersonengruppen

Die Abbildung 29 zeigt die prozentuale Verteilung der intendierten Interaktionen in den einzelnen Versuchspersonengruppen. Die intendierte Interaktion wurde durch die Annotation der Videos ermittelt und unterscheidet sich im vorliegenden Versuche durch die *Dauer* und den *aufgewendeten Druck*, woraus die vier Interaktionstypen *Short tap*, *Long tap*, *Short press* und *Long press* resultieren.

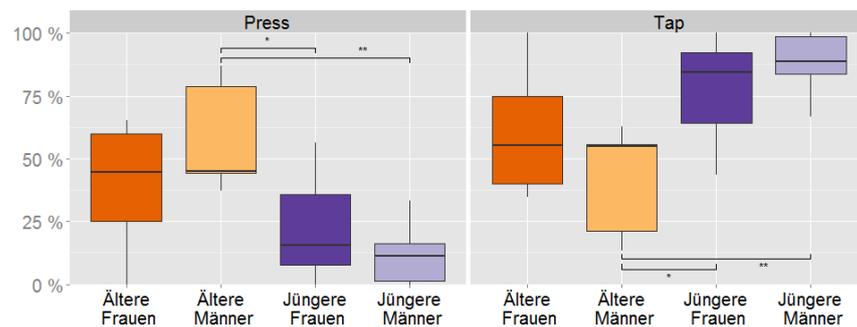
Wie in Abbildung 29 dargestellt, ist bei der Gruppe der jüngeren Versuchspersonen der *Short tap* – die einzig richtige Interaktion für die verwendete App auf dem Smartphone – die dominierende intendierte Interaktion. Für den Anteil der Verwendung von *Short tap* zeigt die einfaktorielle ANOVA signifikante Unterschiede ($F_{df(3,22)} = 5,481$; $p = 0,006$) zwischen den Versuchspersonengruppen. Der paarweise Post-hoc-T-Test führt zu dem Ergebnis, dass signifikante Unterschiede zwischen jüngeren Frauen und älteren Männern ($p < 0,05$) bzw. hoch signifikante Unterschiede zwischen jüngeren Männern und älteren Männern ($p < 0,001$) für die Verwendung von *Short tap* als intendierte Interaktion existieren, wobei die Effektstärken für beide Vergleich groß ($|g| > 0,5$) sind.

Für die Häufigkeiten von *Short press* findet sich ein ähnliches Bild. Wieder lassen sich zwischen den Versuchspersonengruppen durch die ANOVA signifikante Unterschiede ($F_{df(3,22)} = 4,515$; $p = 0,013$) finden. Der Post-hoc-T-Test zeigt signifikante Unterschiede in der Verwendung von *Short press* zwischen den Versuchspersonengruppen der älteren Männer und jüngeren Männer ($p < 0,05$). Alle anderen Paarvergleiche sind die Unterschiede nicht signifikant.

Die *Long tap* als auch die *Long press*-Interaktionen werden in beiden Gruppen nur selten verwendet. Wie in Abbildung 29 zu sehen, nutzen ältere Versuchspersonen diese Interaktionstypen häufiger als jüngere Versuchspersonen, der Unterschied ist allerdings weder für *Long tap* noch für *Long press* signifikant.



(a) Long vs. Short



(b) Press vs. Tap

Abbildung 30: Häufigkeiten in den Dimensionen Zeit und aufgewendeter Druck

Für die intendierte Interaktion kann zusammenfassend festgehalten werden: Die älteren Versuchspersonen nutzen häufiger als jüngere Versuchspersonen andere Interaktionen als *Short tap*. Worin der Unterschied zwischen älteren und jüngeren Versuchspersonen genau besteht, zeigen die beiden nachfolgenden Darstellungen Abbildung 30a und Abbildung 30b. Wie schon die intendierten Interaktionen, wurden auch die Unterscheidungen zwischen *kurz/lang* und *mit Druck/ohne Druck* anhand der Videos annotiert.

Wie in der Abbildung 30a zu sehen ist, sind kurze Interaktionen in allen Gruppen die dominierenden Interaktionsformen. Um Unterschiede zwischen den Versuchspersonengruppen zu finden, wurde erst überprüft, ob die Verteilung der Varianzen (Homoskedastizität) zwischen den Gruppen gleich ist. Bei der Prüfung der Homoskedastizität zeigen sich signifikante Unterschiede ($F_{df(3,22)} = 5,131; p = 0,008$) zwischen den Versuchspersonengruppen, weshalb als Alternative zur parametrischen ANOVA der Kruskal-Wallis Rangsummentest genutzt wurde, welcher knapp signifikante Unterschiede ($\chi^2 = 8,245; df = 3; p = 0,041$) zwischen den Gruppen fand. Der Post-hoc-Test, diesmal bedingt durch die fehlende Homoskedastizität ein paarweiser Wilcoxon-Test mit Bonferroni-Holm-Korrektur, zeigte keine signifikanten Unterschiede zwischen den Gruppen, wobei der Unterschied zwischen

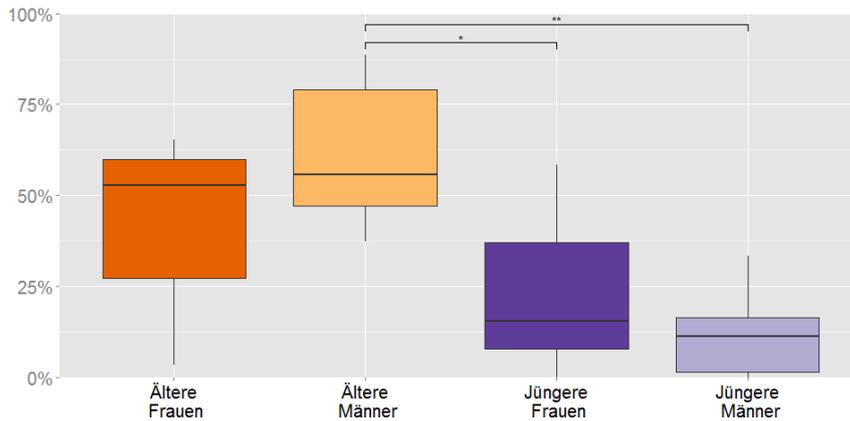


Abbildung 31: Anteile falsch intendierter Interaktionen (Mistakes)

jungen und älteren Männern nur knapp nicht signifikant ($p = 0.055$) war. Für die Dimension *Zeit* kann somit festgehalten werden, dass keine Unterschiede zwischen den Versuchspersonengruppen existieren.

In der Abbildung 30b sind die Verteilungen der Häufigkeiten für die Dimension *aufgewendeter Druck* dargestellt. Wie der Abbildung zu entnehmen ist unterscheiden sich die Gruppen hinsichtlich der Häufigkeit des aufgewendeten Drucks. Die durchgeführte parametrische ANOVA zeigt hoch signifikante Unterschiede zwischen den Versuchspersonengruppen ($F_{df(3,22)} = 5,438$; $p = 0,006$). Der Post-hoc-T-Test zeigt signifikante Unterschiede mit großen Effektstärken zwischen älteren Männern und jüngeren Frauen ($p = 0,047$; $|g| = 1,464$) bzw. älteren Männern und jüngeren Männern ($p = 0,007$; $|g| = 2,523$). Die Unterschiede in der Dimension aufgewendeter Druck legen die Vermutung nahe, dass die Teilnehmer der älteren männlichen Versuchspersonengruppe häufiger ein unpassendes mentales Modell von der Funktionsweise des verwendeten kapazitiven Touchscreens haben, da die Personen dieser Gruppe überdurchschnittlich oft Druck aufwenden, wenn sie mit dem Smartphone interagieren.

Wie sich die falsch intendierten Interaktionen auf die Interaktion mit dem System auswirken, wird im nachfolgenden Unterabschnitt näher erläutert.

4.1.7.3 Fehlerhafte Interaktionen

Abgeleitet aus den intendierten Interaktionen kann für die Versuchspersonengruppen bestimmt werden, wie hoch der Anteil der falsch intendierten Interaktionen (Mistakes) ist. Eine Interaktion wird, für den vorliegenden Versuch, als falsch bzw. als Mistake klassifiziert, wenn sie kein *Short tap* ist.

Die Abbildung 31 stellt den Anteil der Mistakes an den Gesamtinteraktionen in den einzelnen Versuchspersonengruppen dar. Die durchgeführte ANOVA bestätigt, dass zwischen den Gruppen signifikante

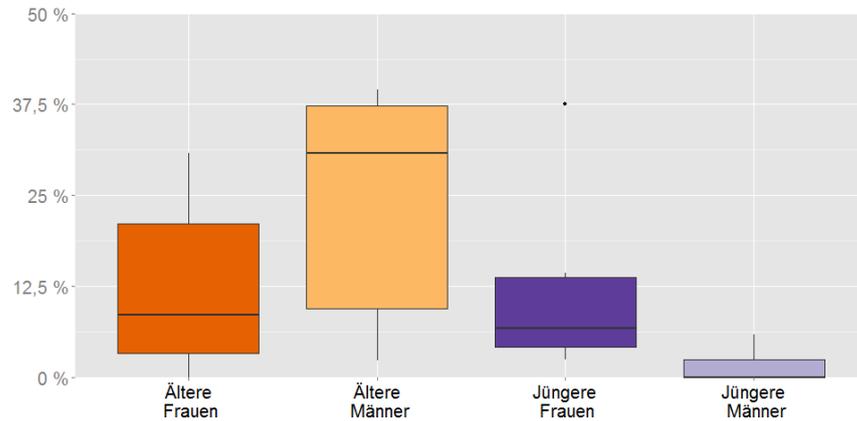
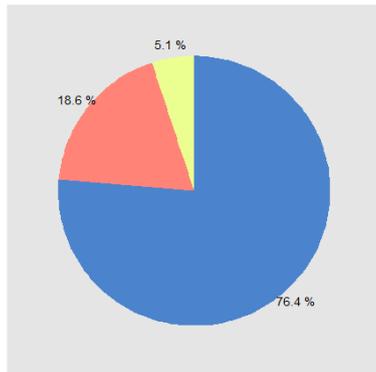


Abbildung 32: Anteile der vom System nicht verarbeiteten Interaktionen

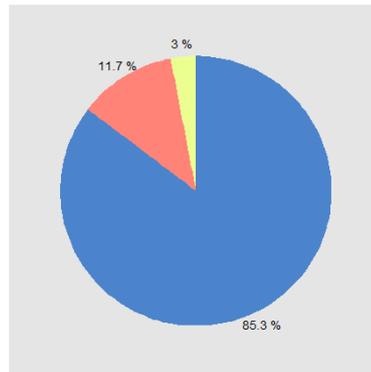
Unterschiede ($F_{df(3,22)} = 6,321; p = 0,003$) bestehen. Der paarweise Post-hoc-T-Test mit sequenzieller Bonferroni- Holm-Korrektur bestätigt die Vermutung, dass signifikante Unterschiede zwischen älteren Männern und jüngeren Männern ($p = 0,003$) bzw. älteren Männern und jüngeren Frauen ($p = 0,028$) bestehen. Für die anderen Paarvergleiche werden keine signifikanten Unterschiede gefunden, wobei der Paarvergleich zwischen jüngeren Männern und älteren Frauen nur knapp nicht signifikant ($p = 0,051$) ist. Als nächstes wird betrachtet, wie hoch der Anteil der nicht vom System verarbeiteten Interaktionen ist.

In Abbildung 32 ist der Anteil der Interaktionen dargestellt, der vom Smartphone bzw. von der App nicht erkannt werden konnte. Da die Prüfung der Homoskedastizität signifikante Unterschiede ($F_{df(3,22)} = 5,723; p = 0,005$) in der Varianz zwischen den Versuchspersonengruppen zeigt, wurde der Kruskal-Wallis Rangsummentest genutzt, um zu testen ob signifikante Unterschiede zwischen den Gruppen bestehen. Wie bei den falsch intendierten Interaktionen zeigt der Vergleich signifikante Unterschiede ($\chi^2 = 10,473; df = 3; p = 0,015$) zwischen den Versuchspersonengruppen, allerdings kann beim paarweisen Vergleich, unter Verwendung des korrigierten Wilcoxon-Tests, kein signifikanter Unterschied zwischen zwei Gruppen gefunden werden. Knapp nicht signifikant sind die Unterschiede zwischen älteren und jüngeren Männern ($p = 0,082$) bzw. zwischen jüngeren Männern und älteren Frauen ($p = 0,059$). Einen starken Zusammenhang zwischen der Verwendung von falschen Interaktionstypen – den sogenannten Mistakes – und dem Auftreten von Erkennungsfehler zeigt der hohe Wert für die Korrelation der Werte (Spearman's ρ von 0,686). In den nachfolgenden Abbildungen wird deshalb aufgeschlüsselt, wieso eine Interaktion nicht erfolgreich war.

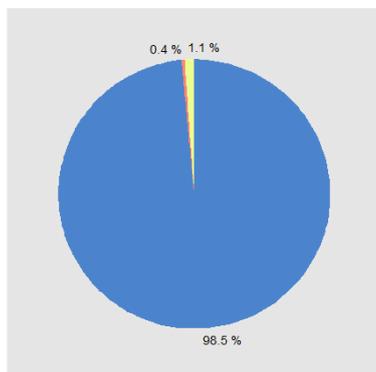
Die Abbildung 33 zeigt für die einzelnen Versuchspersonengruppen, wie häufig die Interaktionen erfolgreich (blau) bzw. nicht erfolgreich waren. War eine Interaktion fehlerbehaftet, so ist dargestellt warum sie nicht erfolgreich war (gelb = *Slip*, rot = Fehler nach *Mistake*).



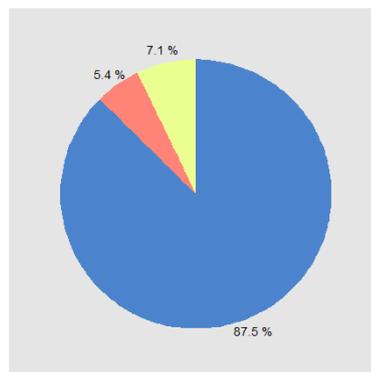
(a) Ältere Männer



(b) Ältere Frauen



(c) Jüngere Männer



(d) Jüngere Frauen

Abbildung 33: Anteile von erfolgreichen und nicht erfolgreichen Interaktionen

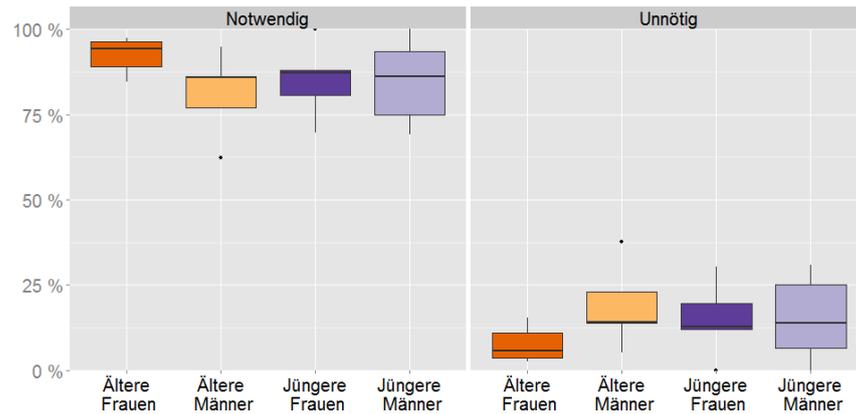


Abbildung 34: Notwendige vs. unnötige Interaktionen

Der Vergleich der Mittelwerte, für nicht erfolgreiche Interaktionen die nach einem Mistake auftraten, brachte signifikante Unterschiede zwischen den Versuchspersonengruppen ($\chi^2 = 10,877$; $df = 3$; $p = 0,012$). Der durchgeführte Post-hoc-Test führte zu dem Ergebnis, dass sich die Anteile der durch Mistakes bedingten nicht erfolgreichen Interaktionen zwischen älteren männlichen Versuchspersonen und jüngeren männlichen Versuchspersonen signifikant unterscheiden ($p = 0,025$).

Für die Anteile der durch Slips nicht erfolgreichen Interaktionen zeigte sich ein ähnliches Bild. Wieder wurden signifikante Unterschiede ($F_{df(3,22)} = 3,367$; $p = 0,037$) zwischen den Versuchspersonengruppen gefunden. Der Post-hoc-T-Test brachtet das Ergebnis, dass in diesem Fall der signifikante Unterschied ($p = 0,041$) zwischen den beiden jüngeren Versuchspersonengruppen, demnach zwischen jüngeren Frauen und jüngeren Männern, besteht.

Neben den fehlerhaften Interaktionen führten auch unnötige Interaktionen zu Wiederholungen von Interaktionen. Im nachfolgenden Unterabschnitt werden die Anzahl der unnötigen Interaktionen und deren vermutete Gründe dargestellt.

4.1.7.4 Unnötige Interaktionen

Die Abbildung 34 zeigt das Verhältnis von nötigen und unnötigen Interaktionen in den verschiedenen Versuchspersonengruppen. Unnötige Interaktionen sind wiederholte Aktionen einer Versuchsperson, die ausgeführt werden, obwohl die Aufgabe bereits erfüllt wurde. Unnötige Interaktionen resultierten im durchgeführten Experiment oftmals aus dem nicht wahrgenommenen Feedback des Systems oder aus Systemfehlern. Bei dem Vergleich der unnötigen Interaktionen zeigt die einfaktorielle Varianzanalyse keine Unterschiede zwischen den Gruppen. Es gibt also keine Gruppe die mehr unnötige Interaktionen aufgewendet hätte als eine andere Versuchspersonengruppen.

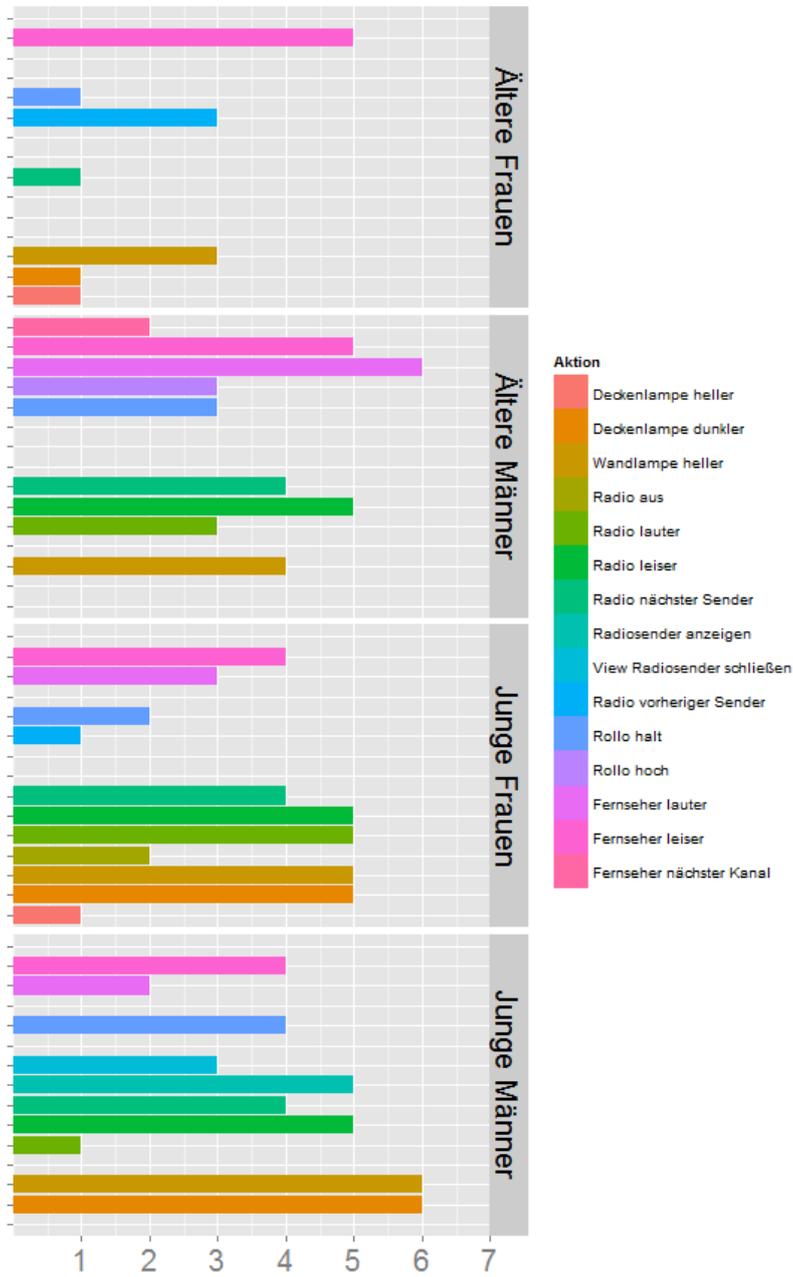


Abbildung 35: Verteilung der unnötigen Interaktionen

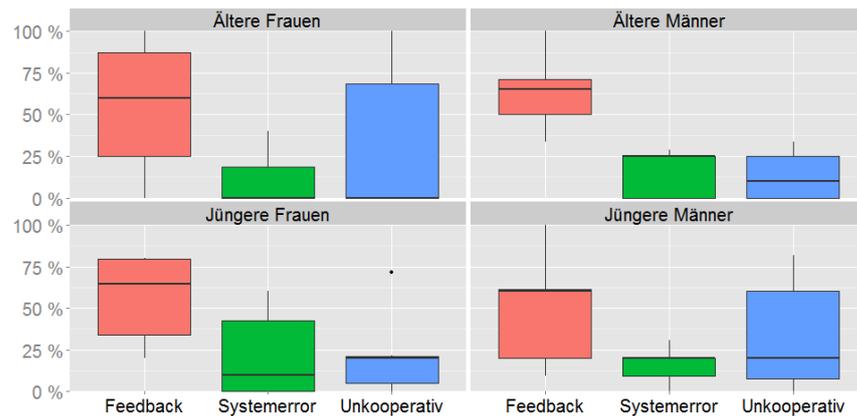


Abbildung 36: Anteile der Gründe für eine unnötige Interaktion

In Abbildung 35 sind die Häufigkeiten für Aktionen aufgelistet, die von mehr als zwei Personen wiederholt wurden, obwohl die vorhergehende Interaktion erfolgreich war. Aktionen, bei denen ein Gerät lauter oder leiser bzw. heller oder dunkler gemacht werden musste, wurden häufiger wiederholt, da es nur geringe Unterschiede zwischen den Lautstärken/Helligkeiten gab, so dass das Systemfeedback von den Versuchspersonen oftmals nicht wahrgenommen wurde. Andere Aktionen, wie der Wechsel des Radiosenders oder die Rollladenbedienung, waren oftmals durch Systemfehler beeinträchtigt; So waren die (Internet-) Radiosender stellenweise nicht erreichbar oder einzelne Rollläden stoppten nicht, obwohl die Aktion Rollläden anhalten vom System verarbeitet wurde.

Die Abbildung 36 stellt die annotierten Gründe für die unnötigen Interaktionen in den unterschiedlichen Versuchspersonen dar. Eine einfaktorische Varianzanalyse ergab keine Unterschiede zwischen den Versuchspersonengruppen. Bei dem Vergleich der Gründe (über alle Versuchspersonengruppen) zeigten sich hochsignifikante Unterschiede ($F_{df(3,22)} = 15,64; p < 0,001$) zwischen den Häufigkeiten. Der paarweise Wilcoxon-Test mit Bonferroni-Holm-Korrektur zeigt, dass unnötige Interaktionen zum überwiegenden Teil aufgrund des nicht wahrgenommenen Feedbacks geschehen.

4.1.8 Zusammenfassung und Ableitungen zur Modellbildung

Wie in den vorherigen Abschnitt dargestellt, benötigen die Versuchspersonen der älteren Versuchspersonengruppe mehr Interaktionsschritte, um die gestellten Aufgaben zu lösen. Nach Kang und Yoon [124] ist der Unterschied zwischen älteren und jüngeren Versuchspersonen, hinsichtlich der benötigten Interaktionsschritte, nicht durch Hintergrundwissen oder Vorerfahrungen beeinflusst. Demgegenüber bestätigt die hohe Korrelation zwischen fehlerhaft intendierten Interaktionen (Mi-

stakes) und vom System nicht verarbeiteten Eingaben die Annahme, dass falsche mentale Modelle zu höheren Fehlerquoten führen. Da die Versuchspersonen aus der Gruppe der älteren Versuchspersonen signifikant häufiger falsche Interaktionen anwenden, was nach Hanson [87] häufig auf die fehlende Erfahrungen im Umgang mit modernen Eingabegeräten zurückzuführen ist, kann davon ausgegangen werden, dass die älteren Teilnehmer häufiger falsche mentale Modelle davon hatten, wie ein Smartphone mit einem kapazitiven Touchscreen bedient werden muss.

- Für das Modell zur Vorhersage von Interaktionsfehlern ist die Integration von fehlerhaften intendierten Interaktionen aufgrund von unpassenden mentalen Modellen des zu bedienenden Gerätes sehr wichtig.

1. Ableitung

Ein weiteres Problem welches nicht vernachlässigt werden darf, ist das Problem der Nicht-Wahrnehmung des Feedbacks. Ein fehlendes oder fehlerbehaftetes Feedback führte häufig zu unnötigen Wiederholungen. Da die älteren Versuchspersonen häufiger unpassenden Interaktionstypen auf den Touchscreen angewendet haben und diese Interaktionen häufiger zu Fehlern in der Erkennung durch das System führten, ist fehlerhaftes oder fehlendes Feedback ein Problem.

- Fehlendes Feedback führt im Allgemeinen zu unnötigen Wiederholungen. Häufige Wiederholungen führen zu mehr falschen Interaktionen, welche zu mehr nicht erfolgreichen Interaktionen führen.

2. Ableitung

4.2 ERWEITERUNG DER MEMO-WERKBANK

In diesem Abschnitt werden die Erweiterungen an der MeMo-Werkbank vorgestellt, welche umgesetzt wurden, um Fehler durch falsche bzw. unpassende mentale Modelle zu simulieren. Der erste Unterabschnitt beschreibt das Konzept der Generischen Interaktionen und das Konzept der Device Models. Beide Konzepte bilden die Grundlage dafür, dass Fehler durch falsche oder unpassende mentale Modelle ermöglicht werden. Der anschließende Unterabschnitt beschreibt die Integration von Systemfeedback und neuen Nutzerattributen in die Werkbank. Wie in der Zusammenfassung des empirischen Versuchs dargestellt, ist die Integration von Feedback in die Werkbank sehr wichtig, da ein fehlerhaftes oder fehlendes Feedback zu wiederholten Interaktionen führen kann, welche wiederum fehlerbehaftet sein können. In den nachfolgenden drei Unterabschnitten werden die neuen Module zur Wahrnehmung, Verarbeitung und Ausführung des Nutzermodells beschrieben, welche notwendig sind, um Interaktionsfehler aufgrund falscher oder unpassender mentaler Modell zu simulieren. Den Abschluss des Abschnitts bilden die Beschreibung des neuen Modules

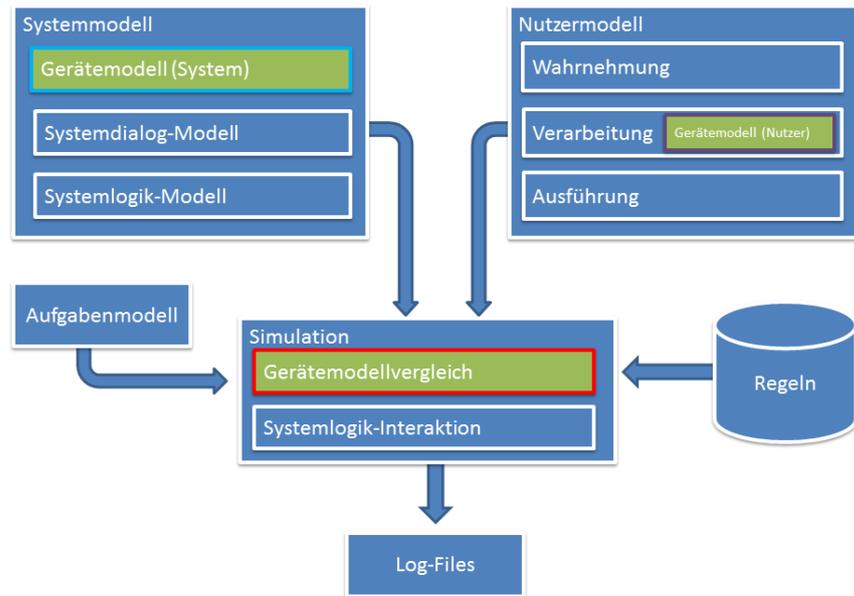


Abbildung 37: Aufbau der MeMo-Werkbank

zur Interpretation der Nutzermodellinteraktion und die Beschreibung eines neuen Moduls zum Export der Simulationsdaten.

4.2.1 Generische Interaktionen und Device Models

In Abbildung 37 ist der schematische Aufbau der MeMo-Werkbank dargestellt. Um Fehler zu simulieren, die auf ein unpassendes oder falsches mentales Modell zurückzuführen sind, wurden sogenannte Generische Interaktionen implementiert. Zur Integration der generischen Interaktionen mussten die mit grün dargestellten Module in die bestehende Werkbank neu entwickelt werden.

War es in den früheren Implementierungen nur möglich Systemmodelle zu entwerfen, welche per Maus oder Sprache bedient werden konnten, so ermöglicht die Erweiterung der Werkbank durch generische Interaktionen nun, diverse Eingabegeräte zu simulieren. Generische Interaktionen formen hierbei eine neue Schicht in der Simulation der MeMo-Werkbank (mit einem roten Rahmen markiert) und trennen das Gerätemodell des Systems (Hellblauer Rahmen) vom der Systemlogik (Verknüpfung der Systemzustände). In früheren Implementierungen waren Systemzustände immer über Mausinteraktionen oder Spracheingaben miteinander verknüpft, wobei die Mausinteraktionen maustypische Eigenschaften (Links-Klick, Rechts-Klick, usw.) hatten. In der aktuellen Werkbank sind die Systemzustände durch eigenschaftslose Interaktionen (Generische Interaktionen) verknüpft. Eine generische Interaktion hat einen Namen und einen Interaktionstyp.

Ein Interaktionstyp ist ein Hilfskonstrukt ohne eigene Eigenschaften, welches nur dafür genutzt wird verschiedene Interaktionen auf einem Interaktionsobjekt voneinander zu unterscheiden.

Ein Beispiel soll Begriffe Generische Interaktionen und Interaktionstyp veranschaulichen und deren Verhältnis zueinander erklären.

Zwei Systemzustände A und B sind über einen Button miteinander verknüpft. In älteren Werkbankimplementierungen wurden die beiden Zustände z. B. über einen Links-Klick miteinander verbunden. Im Kontext der generischen Interaktionen werden die beiden Zustände nun von einer generischen Interaktion mit dem Namen *A_nach_B* und vom Typ *GENERIC_01* miteinander verbunden. Die generische Interaktion alleine besagt nicht, wie der Button bedient werden muss, damit das System vom Zustand A in den Zustand B wechseln kann. Erst wenn vom Modellierer definiert wird, dass das Eingabegerät (Gerätemodell des Systems) ein kapazitiver Touchscreen ist, wird definiert, wie der Button bedient werden muss. Im Fall des kapazitiven Touchscreens als Systemeingabegerät muss das Nutzermodell einen *Short tap* auf den Button anwenden, um die Transition auszuführen, da für das Device Model *Kapazitiver Touchscreen* die Interaktion *Short tap* mit dem Interaktionstyp *GENERIC_01* verbunden ist. Wechselt der Modellierer das Eingabegerät, z. B. auf einen resistiven Touchscreen, dann ändert sich auch die Art wie der Button bedient werden muss; jetzt muss das Nutzermodell einen *Short press* ausführen, damit die Transition erfolgen kann. Insgesamt existieren in der MeMo-Werkbank 20 verschiedene generische Interaktionstypen (*GENERIC_01*, *GENERIC_02*, ..., *GENERIC_20*) die zur Definition von Systemzustandsübergängen verwendet werden können. Hieraus resultiert, dass mit der aktuellen Werkbank nur Eingabegeräte simuliert werden können, welche höchstens 20 unterscheidbare Eingabeinteraktionen bieten.

In der erweiterten Werkbank sind 5 verschiedene Gerätemodelle implementiert. Neben zwei Maus-Modellen (3-Button-Maus, Maus mit Scrollrad) und zwei Touchscreen-Modellen (resistiver Touchscreen, kapazitiver Touchscreen) wurde noch ein Fernbedienungsmodell hinzugefügt. Die genannten Device Models können sowohl als Gerätemodell des Systems als auch als Gerätemodell des Nutzermodells (mit einem lila Rahmen markiert) in der Werkbank genutzt werden.

Bei der Verwendung eines Device Models als mentales Gerätemodell, als Device Mental Model (DMM), wird es zum Nutzerwissen hinzugefügt. Das Hinzufügen zum Nutzerwissen geschieht im Nutzergruppeneditor (vgl. Abschnitt 2.1.5). Das nachfolgende Beispiel soll erläutern, was passiert, wenn ein DMM zur Beschreibung der Nutzermodellgruppe hinzugefügt wird.

Wird das Gerätemodell *Kapazitiver Touchscreen* zum Nutzerwissen hinzugefügt, so bedeutet dies, dass das Nutzermodell weiß, wie es einen kapazitiven Touchscreen bedienen muss. Weiterhin weiß das Nutzermodell hierdurch, welche Interaktionstypen möglich sind.

Da das neue MeMo-Nutzermodell mehr als ein mentales Device Modell haben kann, kann das Nutzermodell während der Interaktion mit dem System das verwendete mentale Gerätemodell wechseln. Durch die Integration von Device Models in die Werkbank ist eine wichtige Grundlage geschaffen, dass Interaktionsfehler simuliert werden können, die auf fehlerhaften mentalen Gerätemodellen basieren.

Eine letzte Stelle, an welcher die Device Models verwendet werden, ist im Simulator selbst. Hier wird überprüft, ob das Systemgeräte-Modell und das mentale Gerätemodell vom selben Typ sind oder ob sie sich unterscheiden. Sind die verwendeten Gerätemodell vom selben Typ wird die Interaktion ohne Interpretation vom System verarbeitet, unterscheiden sich die beiden Gerätemodelle, so muss ein Mapping der benutzten Interaktion auf den Raum der vom System verstehbaren Interaktionen geschehen. Zur Berechnung ob eine Nutzermodell-Interaktion vom System verstanden werden kann, wird ein Modul zur Verarbeitung der Nutzereingabe (siehe: Unterabschnitt 4.2.6) verwendet.

4.2.2 Systemfeedback und neue Nutzerattribute

Wie durch den Versuch gezeigt, kann fehlendes oder fehlerhaftes Feedback dazu führen, dass Eingaben wiederholt werden, was die Anzahl der benötigten Eingaben erhöht, aber auch zu mehr Fehlern in der Bedienung führen kann. Ein Systemfeedback war in der bisherigen Werkbank nicht implementiert, dabei ist (System-)Feedback ein kritischer Faktor im Bereich der Mensch-Maschine-Interaktion, da es dazu dient dem Nutzer zu vermitteln was gerade geschieht [173, S. 27]. Schlechtes Feedback ist besonders kritisch, wenn schlecht designte Objekte so konstruiert sind, dass falsche mentale Modelle zu Fehlern führen können [173, S. 42].

Nach Vasilyeva et al. [238] kann Feedback hinsichtlich der Dimensionen: Grad der Anpassung, Kontext der Anpassung, Zeit des Auftretens, Funktion, Komplexität, Form der Präsentation, Grad der Information, Intension des Feedbacks, Ziel des Feedbacks und Abdeckung des Fortschritts klassifiziert werden. Diese Klassifikation ist sehr detailliert, allerdings unpraktisch für die Simulation, da vom Modellierer viele Informationen zu spezifizieren sind, welche im Falle einer Fehleinschätzung durch den Modellierer zu unrealistischen Simulationsergebnissen führen können. Die neue Implementierung von Feedback in der MeMo-Werkbank ist an die Klassifikation von Dix et al. [54] angelehnt.

Nach der Definition von Dix und Kollegen kann zwischen *physischem Feedback*, *systemgesteuertem physischem Feedback*, *systemgesteuertem virtuellem Feedback* und *indirektem physischem Feedback* unterschieden werden. Physisches Feedback beschreibt die direkte physikalische Wahrnehmung einer Person wenn diese mit einem System interagiert

(z. B. wenn die Person fühlt wie sie einen Knopf drückt). Systemgesteuertes physisches Feedback ist zeitnahes Feedback, welche durch das System ausgelöst wird und das Eingabegeräte beeinflusst (z. B. das Vibrieren des Telefons bei der Berührung des Touchscreens). In die Kategorie systemgesteuertes virtuelles Feedback fallen z. B. Anzeigen auf dem Bildschirm oder Status-LEDs. Indirektes physikalisches Feedback ist Feedback in der Umgebung (z. B. das Einschalten der Deckenbeleuchtung). Da auch die Klassifikation von Dix noch sehr detailliert ist, wurde für die MeMo-Werkbank eine weitere Vereinfachung angenommen und es wurden die Kategorien *kein Feedback*, *indirektes Feedback* und *direktes Feedback* abgeleitet.

Neben dem Typ des Feedbacks muss auch die Qualität desselben in die Simulation einfließen. Da die Qualität eine subjektive Einschätzung eines Ereignisses ist und dem Autor keine Modelle bekannt sind, die in simulierten Umgebungen die Feedbackqualität vorhersagen können, muss für die Modellierung wiederum eine Vereinfachung getroffen werden. Die Vereinfachung zielt auf die Beschreibung der Intensität des Feedbacks ab. Hierfür muss der Modellierer für das definierte Feedback einschätzen wie hoch der Level des Feedbacks ist (angelehnt an Vasilyeva et al. [238]). Die Angabe des Levels erfolgt auf einer 5-stufigen Skala von *sehr niedrig* bis *sehr hoch*.

Neben der Implementierung von Systemfeedback sind für die Berechnung der mentalen Modelle (siehe Abschnitt 4.2.4) auch neue bzw. erweiterte Nutzerattribute nötig. Die neuen Nutzerattribute basieren größtenteils auf den Items der TA und ICT Fragebögen. So wurde für jedes Item der Fragebögen – insgesamt 49 Items (19 TA und 30 ICT) – ein entsprechendes Attribut in der MeMo-Werkbank angelegt. Weiterhin wurde eine Unterscheidung zwischen männlichen und weiblichen Nutzermodellen durch die Einführung des Attributes Geschlecht ermöglicht. Die Unterscheidung hinsichtlich des Geschlechtes ist notwendig, da die genannten Nutzerattribute geschlechtsspezifische Ausprägungen haben (siehe Abschnitt 4.1.7.1).

4.2.3 Modul zur Wahrnehmung der Benutzerschnittstelle

Da die Wahrnehmung der Systemreaktion einen entscheidenden Anteil daran hat, ob eine Aufgabe als abgeschlossen bewertet wird oder ob die Aufgabe wiederholt werden muss, wird ein neues Wahrnehmungsmodul benötigt. Dieses Wahrnehmungsmodul (engl. Perception-Modul | PEM) muss in der Lage sein, das modellierte Systemfeedback auszuwerten und auf Basis der Auswertung zu entscheiden, ob das Feedback vom Nutzermodell wahrgenommen wird oder nicht. Das neu entwickelte PEM leitet sich vom Default-Perception-Modul ab. Es umfasst alle Funktionen des zugrundeliegenden Moduls und erweitert dieses Modul durch einen Algorithmus zur Wahrnehmung des Feedbacks.

Gruppe	Level des Feedbacks					
	kein	sehr gering	gering	normal	hoch	sehr hoch
Ält. Männer	1	0,594	0	0	0	0
Ält. Frauen	1	0,357	0	0	0	0
Jüng. Männer	1	0,410	0	0	0	0
Jüng. Frauen	1	0,369	0	0	0	0

Tabelle 24: Wahrscheinlichkeiten: Feedback nicht wahrgenommen

Zur Auswertung des Feedbacks überprüft der Algorithmus zuerst, ob vorher eine Interaktion mit dem System stattgefunden hat. Dies ist notwendig, da mit der MeMo-Werkbank bisher nur Systeme simuliert werden können, bei denen das Nutzermodell die Initiative übernimmt.

Ist die Nutzermodell-Aktion nicht die erste Interaktion mit dem Systemmodell, wird vom Algorithmus überprüft, ob vom Modellierer ein Feedback definiert wurde. Hat der Modellierer den Systemübergang zwischen den Systemzuständen mit *kein Feedback* definiert, so wird das Feedback, vom Wahrnehmungsmodul, als nicht wahrnehmbar markiert. Ist das Feedback von einem anderen Typ (*direktes Feedback* oder *indirektes Feedback*), so wird anhand des Feedback-Levels, der Altersgruppe und des Geschlechts berechnet, ob das Wahrnehmungsmodul das Feedback wahrnimmt.

In Tabelle 24 sind die Wahrscheinlichkeiten dargestellt, dass ein Nutzermodell das Feedback nicht wahrnimmt und deshalb die Interaktion wiederholt. Die in der Tabelle angegebenen Wahrscheinlichkeiten ergeben sich aus den empirischen Daten des Versuchs, für die Level *gering*, *normal* und *hoch* wurden keine empirischen Daten erhoben, weshalb die Wahrscheinlichkeit, dass das Feedback nicht wahrgenommen wird, mit Null definiert ist. Für die Wahrnehmung des Feedbacks bedeutet dies: Hat das System kein Feedback, so wird das Nutzermodell auch kein Feedback wahrnehmen. Hat das System ein geringes bis sehr hohes Feedback, so wird dieses Feedback vom Nutzermodell erkannt. Für den Fall, dass das System ein sehr geringes Feedback hat, wird ein Algorithmus verwendet. Um zu entscheiden ob ein Feedback wahrgenommen wird oder nicht, wird mit einem Zufallsalgorithmus eine Zahl zwischen 0 und 1 generiert. Ist die Zahl für die entsprechende Nutzergruppe kleiner als der in Tabelle 24 dargestellte Schwellwert, so wird das Feedback vom Nutzermodell nicht erkannt.

Nachdem bestimmt wurde, ob das Feedback erkannt wurde, werden die in der Benutzerschnittstelle wahrgenommenen Interaktionsobjekte (Button, Links und ähnliches) an das Modul zur Verarbeitung der

wahrgenommenen Ereignisse transferiert, wo diese ausgewertet werden.

4.2.4 Modul zur Verarbeitung von wahrgenommenen Ereignissen

Die existierenden verschiedenen Module zur Verarbeitung der wahrgenommenen Ereignisse berechnen, in Abhängigkeit von den Eigenschaften der möglichen Interaktionsoptionen und in Abhängigkeit vom Aufgabenwissen des Nutzermodells, die Wahrscheinlichkeiten für die Interaktionsoptionen. Basierend auf den Wahrscheinlichkeiten wird eine Interaktion aus den angebotenen Optionen ausgewählt. Die Auswahl der Interaktion erfolgt allerdings bislang, ohne dass das Eingabegerät oder das mentale Modell des Nutzermodells einen Einfluss auf die Auswahl hätte.

Das neue Modul zur Verarbeitung von wahrgenommenen Ereignissen (engl. Processing-Modul | PROM) soll nicht nur entscheiden, welche Interaktion der angebotenen Interaktionsoptionen ausgeführt wird, sondern auch in der Lage sein, das Eingabegerät in die Entscheidung mit einzubeziehen. Hierfür musste der Prozess wie eine Interaktion ausgewählt wird überarbeitet werden.

Wie bereits dargestellt, werden in den anderen PROMs nur die Interaktionen betrachtet und diese dann in Abhängigkeit vom Aufgabenwissen und den Interaktionsoptioneigenschaften verarbeitet. Im neuen PROM wird in einem ersten Schritt entschieden, ob eine Aufgabe bzw. eine Interaktion wiederholt wird oder nicht. Daran anschließend wird entschieden, welches mentale Modell Anwendung finden wird. Ist das mentale Modell gewählt, werden die Eigenschaften der verschiedenen Interaktionsobjekte analysiert, um die Auswahlwahrscheinlichkeiten für die Interaktionsobjekte zu berechnen. Sind die Wahrscheinlichkeiten für die Interaktionsobjekte berechnet wird durch einen Zufallsalgorithmus ein Interaktionsobjekt unter Beachtung der Wahrscheinlichkeiten ausgewählt. Ist das Interaktionsobjekt ausgewählt, wird anhand des mentalen Modells und der vorhergehenden Interaktionen die *intendierte Interaktion* (Intended Interaction) berechnet. Im Folgenden werden die einzelnen Schritte der Interaktionsauswahl detaillierter erläutert.

Die Entscheidung, ob eine Aufgabe bzw. ein Interaktionsschritt wiederholt wird, ist in der aktuellen Implementierung sehr einfach. Da weder unkooperatives Verhalten, noch Systemfehler simuliert werden sollen, hat allein das Systemfeedback bzw. das Fehlen des Systemfeedbacks einen Einfluss darauf, ob ein eigentlich abgeschlossener Interaktionsschritt wiederholt wird. Für das Nutzermodell wurde festgelegt, dass das Nutzermodell den letzten Interaktionsschritt wiederholt, wenn das Systemfeedback fehlte oder nicht wahrgenommen wurde.

Um zu entscheiden, welches mentale Gerätemodell vom Nutzermodell für die intendierte Interaktion genutzt werden soll, muss berechnet werden, mit welcher Wahrscheinlichkeit sich ein Nutzermodell für ein Gerätemodell entscheiden wird. Um die Wahrscheinlichkeiten für die möglichen Device Models – im vorliegenden Fall, ein kapazitiver Touchscreen, ein resistiver Touchscreen oder eine Fernbedienung – zu berechnen, werden insgesamt 6 allgemeine lineare Modelle (General linear models | GLMs) genutzt, welche mit den Attributwerten des Nutzermodells initialisiert werden. Die Anzahl der GLMs resultiert aus der Kombination der Anzahl der möglichen mentalen Modelle (kapazitiver Touchscreen/resistiver Touchscreen/Fernbedienung) und ob das Feedback wahrgenommen wurde oder nicht (Wahrnehmung von Feedback/keine Wahrnehmung von Feedback).

Die im Anhang A und in Formel 2 dargestellten allgemeinen linearen Modelle wurden in R unter Zuhilfenahme der Funktion *glmulti* berechnet. Die Funktion *glmulti* bekommt mehrere Vektoren übergeben, wovon ein Vektor den Zielvektor beschreibt und die anderen Vektoren (Variablenvektoren) innerhalb des allgemeinen linearen Modells zur Vorhersage des Zielvektors verwendet werden können. Die Funktion berechnet für alle Kombinationen der Variablenvektoren die passenden allgemeinen linearen Modelle und das gewählte Informationskriterium. Für die vorliegenden Daten wurde das Bayes'sches Informationskriterium (BIC) zur Auswahl des besten allgemeinen linearen Modells verwendet, da dieses Informationskriterium darauf abzielt das optimale Modell zu finden (vgl. Schwarz [211]). Es wird folglich das Modell gewählt, dass mit möglichst wenigen Parametern möglichst viel Varianz in den Daten erklärt.

Bei der Auswahl der Variablenvektoren wurde darauf geachtet, dass diese potenziell einen Einfluss auf die Auswahl des DMMs haben. So wurde z. B. darauf verzichtet die ICT Dimension Design in den Variablenvektorraum aufzunehmen, da davon auszugehen ist, dass die Orientierung an Marken bzw. die Vorliebe für Design keinen Einfluss darauf hat, wie ein Gerät bedient wird. Weiterhin wurde darauf verzichtet einzelne Items der Fragebögen zu verwenden.

Nachfolgend ist die Formel zur Berechnung der Wahrscheinlichkeit der Auswahl des mentalen Modells *Kapazitiver Touchscreen* (P_{KP}) dargestellt. Wenn das Nutzermodell jung ist, werden 0,26541 von der Grundwahrscheinlichkeit (0,87760) abgezogen. Anschließend wird, wenn das Nutzermodell jung ist, der berechnete Wert für die TA-Dimension Begeisterung (BG_{TA}) mit 0,37295 multipliziert und ebenfalls hinzu addiert. Das Vorgehen ist für die nachfolgenden Faktoren äquivalent, wobei für die Erläuterung der Abkürzungen auf Tabelle

Abkürzung	Dimension
A_{ICT}	Technikeinstellung: Angst
BG_{TA}	Technikaffinität: Begeisterung
Exp_{ICT}	Technikeinstellung: Exploration
Int_{ICT}	Technikeinstellung: Interesse
$Komp_{TA}$	Technikaffinität: Kompetenz
Neg_{TA}	Technikaffinität: negative Einstellung zur Technik
Pos_{TA}	Technikaffinität: positive Einstellung zur Technik

Tabelle 25: Abkürzungen im allgemeinen linearen Modell

25 verwiesen wird. Für die anderen 5 Formeln zu Berechnungen der Wahrscheinlichkeiten wird auf den Anhang A verwiesen.

$$\begin{aligned}
P_{KT} = & 0,87760 + (Alt?|0;1) * (-0,26541 + BG_{TA} * 0,37295) + \\
& Neg_{TA} * (-0,19824 + Pos_{TA} * 0,17827) + \\
& Pos_{TA} * (0,42259 + Komp_{TA} * 0,24145 + A_{ICT} * 0,35559 + \\
& Int_{ICT} * -0,21636) + BG_{TA} * (0,09866 + \tag{2} \\
& Komp_{TA} * -0,06430 + A_{ICT} * -0,26446 + \\
& Exp_{ICT} * -0,53592 + (Mann?|1;0) * -0,49623) + \\
& Int_{ICT} * Komp_{TA} * 0,43874
\end{aligned}$$

Das in der Gleichung 2 dargestellte allgemeine lineare Modell reduziert den summierten Fehler der mittleren Standardabweichung zur Vorhersage der Wahrscheinlichkeiten zur Nutzung des Touchscreens von 1,496 bei 20 Freiheitsgraden auf 0,081 bei 10 Freiheitsgraden. Die Anwendung des allgemeinen linearen Modells verringert den Vorhersagefehler demzufolge um den Faktor 18. In Tabelle 26 sind die Werte für die Performanz der anderen allgemeinen linearen Modelle dargestellt.

Sind die GLMs mit den Attributwerten befüllt, liefern diese pro Gerätemodell einen Wahrscheinlichkeitswert. Um zu entscheiden welches mentale Modell für den aktuellen Interaktionsschritt gewählt wird, werden die Wahrscheinlichkeitswerte normalisiert und im Anschluss wird durch einen Zufallsalgorithmus bestimmt welches mentale Gerätemodell genutzt wird.

Wenn das mentale Gerätemodell für den Interaktionsschritt festgelegt ist, wird vom Processing-Modul das Interaktionsobjekt ausgewählt. Die Reihenfolge dieser Schritte beruht auf der Annahme, dass ein Nutzer bereits ein mentales Modell vom Eingabegerät hat, bevor er oder sie ein Bedienelement (Interaktionsobjekt) auf dem Bildschirm

Geräte	Feedback wahrge- nommen	Δ_{pre}	df	Δ_{post}	df
Kapazitiver Touchscreen	ja	1,751	25	0,054	11
Resistiver Touchscreen	ja	1,501	25	0,056	10
Fernbedienung	ja	0,167	25	0,002	13
Kapazitiver Touchscreen	nein	1,496	20	0,081	10
Resistiver Touchscreen	nein	1,304	20	0,038	8
Fernbedienung	nein	0,109	20	0,002	8

Tabelle 26: Performanz der allgemeinen linearen Modelle

auswählt. Für jedes Interaktionsobjekt wird anhand der Eigenschaften des Objektes und der Übereinstimmung mit dem Aufgabenwissen des Nutzermodells eine Wahrscheinlichkeit berechnet. Die Eigenschaften, die zur Bestimmung der Wahrscheinlichkeit genutzt werden, sind Labels, modellierte Input- bzw. Output-Informationen und modellierte Navigationseigenschaften (z. B. weiter, zurück). Stimmt eine Eigenschaft oder mehrere Eigenschaften mit dem Aufgabenwissen des Nutzermodells überein, so wird die Wahrscheinlichkeit für das entsprechende Interaktionsobjekt erhöht. Sind alle Eigenschaften evaluiert, werden die Wahrscheinlichkeiten über alle Interaktionsobjekte normalisiert und ein Zufallsalgorithmus wählt ein Objekt aus.

Nachdem das mentale Gerätemodell und das Interaktionsobjekt bestimmt wurden, kann berechnet werden, welche Interaktion auf das Objekt angewendet wird. Da durch das mentale Modell bereits festgelegt ist, ob bei der Interaktion Druck angewendet wird oder nicht, muss bei der Interaktionsauswahl noch festgelegt werden, ob sich die Interaktion in einer anderen Form (z. B. durch die Dauer) unterscheidet. Hierfür wird durch einen Algorithmus bestimmt, ob ein Nutzermodell eher eine kurze oder eine lange Interaktion durchführen will.

Der Algorithmus zur Bestimmung der Dauer berechnet anhand des gewählten DMMs und nach dem Feedback der letzten Interaktion, ob das Nutzermodell eine kurze bzw. lange Interaktion auf den Touchscreen anwenden wird. Wurde vom Nutzermodell ein Feedback wahrgenommen, so wird die letzte Interaktion als vom System verarbeitet markiert. Wurde kein Feedback registriert, wird die letzte Interaktion als nicht verarbeitet markiert, was die Wahrscheinlichkeiten für kurze bzw. lange Interaktionen verändert. Die Verteilung der Wahrscheinlichkeiten wurde aus den empirischen Daten bestimmt und ist in [Tabelle 27](#) dargestellt.

Mentales Modell	Feedback wahrgenommen	p_{kurz}	p_{lang}
Kapazitiver Touchscreen	ja	0,992	0,008
Kapazitiver Touchscreen	nein	1,0	0,0
Resistiver Touchscreen	ja	0,942	0,058
Resistiver Touchscreen	nein	0,7	0,3
Fernbedienung	ja	0,737	0,263
Fernbedienung	nein	0,588	0,412

Tabelle 27: Wahrscheinlichkeiten für eine kurze bzw. lange Interaktion

Die Berechnung ob das Nutzermodell eine kurze bzw. lange Interaktion ausführt, wird wieder über einen Zufallsalgorithmus bestimmt, der die gewichteten Wahrscheinlichkeiten übergeben bekommt. Nachdem das mentale Modell, das Interaktionsobjekt und die Dauer der Interaktion berechnet sind, kann die resultierende Interaktion ausgewählt und an das Modul zur Manipulation der Benutzerschnittstelle übergeben werden.

4.2.5 Modul zur Manipulation der Benutzerschnittstelle

Neben den Interaktionsfehlern, die aus falschen mentalen Modellen vom Eingabegerät resultieren, existieren Interaktionsfehler, die unbeabsichtigt sind und nicht auf Mistakes zurückzuführen sind. Um diese Interaktionsfehler (Slips) zu simulieren wird ein neues Modul zur Manipulation der Benutzerschnittstelle (engl. Execution-Modul | EM) benötigt.

Das neu geschaffene EM funktioniert bisher nur für Interaktionen auf Touchscreens. Für die Implementierung wird angenommen, dass der Punkt, wo das Nutzermodell den Bildschirm berühren wird, normalverteilt um den Mittelpunkt des intendierten Interaktionsobjektes liegt.

Nach [158] und [219] kann davon ausgegangen werden, dass etwa 4 % aller Interaktionen nicht erfolgreich sind, weil das gewünschte Interaktionsobjekt nicht getroffen wird. Für den empirischen Versuch bestätigen sich diese Zahlen, wenn der Mittelwert über alle Interaktionsobjekte berechnet wird. Werden die Interaktionsobjekte allerdings hinsichtlich der Größe und des Abstandes klassifiziert zeigen sich Unterschiede zwischen den Klassen. So wurden die in Abbildung 26a dargestellten großen Buttons auf der Startseite (TV, Radio, Rollo, Lampen) nur relativ selten (0,5 % der Fälle) nicht getroffen, wohingegen die Fehlerrate für die dicht gestapelten und sehr schmalen Buttons in

Abbildung 26b (z. B. Einschalten, Ausschalten) mit durchschnittlich 5,05 %, höher ausfiel. Für den kleinen Button zum Schließen des Views (Abbildung 26b und Abbildung 26c rechts oben in der Ecke) betrug die Fehlerquote im Durchschnitt 4 %.

Für das EM wurde ein Algorithmus implementiert, welcher basierend auf der Größe des Buttons und dem Abstand des Buttons zu anderen Buttons, mit einer definierten Quote zu einem Fehler (Slip) führt. In der aktuellen Implementierung werden Buttons hinsichtlich der Größe als *groß*, *normal* und *klein* klassifiziert. Diese Vereinfachung ist nötig, da in der aktuellen MeMo-Werkbank die Größe des Bildschirms nicht definiert werden kann, weshalb die Berechnung der Buttongröße nicht möglich ist. Zur Klassifikation der Größe eines Buttons wurde nun die folgende einfache Regel genutzt: Nimmt ein Button in jeder Dimension (x, y) mindestens 25 % des Bildschirms ein, so wird der Button als *groß* klassifiziert. Ist der Button in einer Dimension kleiner als 10 % des Bildschirms, wird der Button als *klein* klassifiziert. In allen anderen Fällen wird der Button als *normal groß* klassifiziert. Für den Abstand der Buttons wurde ein binäres Attribut eingeführt, welches nur definiert, ob ein Abstand zu anderen Interaktionsobjekten besteht oder nicht.

Für den Fall, dass der kleine Button zum Schließen des Views bedient werden soll, ist eine Slip-Rate von 4 % definiert. Dies bedeutet, dass 96 % der Interaktionen den Button treffen und 4 % der Interaktionen nicht. Da ein Button verfehlt werden kann indem darüber, darunter, rechts oder links der Bildschirm berührt wird, muss die Wahrscheinlichkeit, dass ein Button getroffen bzw. verfehlt wird, zwischen den Dimensionen (x und y) aufgeteilt werden. Um im Mittel den Button mit einer 96 % Wahrscheinlichkeit zu treffen, muss der Button in der x- und in der y-Dimension mit 97,98 % ($\sqrt{96} = 97,97959$) Wahrscheinlichkeit in den beiden Dimensionen getroffen werden.

Legt man nun zwei Gauß'sche Glockenkurven über den Button – eine in Richtung der x-Dimension und eine in Richtung der y-Dimension – und definiert, dass die Grenzen des Buttons 97,98 % der Fläche unter diesen Glockenkurven definieren, so kann man die z-Werte für die Konfidenzintervalle berechnen, welche für den genannten Wert jeweils zwischen -2,323 und 2,323 liegt.

Mit Hilfe einer Zufallsfunktion, die als Basis eine Normalverteilung mit dem Mittelwert 0 und der Standardabweichung 1 nutzt, werden zwei Werte für x und y bestimmt. Ein Button gilt als nicht getroffen, wenn einer der zufälligen z-Werte größer als 2,323 oder kleiner als -2,323 ist, für alle anderen Fälle gilt der Button als getroffen.

Liegen die Werte für x und y vor, so kann der Punkt auf dem Touchscreen berechnet werden, auf dem das Nutzermodell den Bildschirm berührt. Hierfür werden als erstes die Faktoren für die Höhe und Breite eines Buttons berechnet. Die nachfolgende Formel 3 zeigt die Berechnung des Höhenfaktors. Die Variable h_B steht in dieser Formel

für die Höhe des Buttons in Pixeln, z_{FR+} steht für den entsprechenden positiven z-Wert der Fehlerrate und f_H ist der Faktor der Höhe. Der konstante Faktor 2 resultiert aus der Annahme der Voraussetzung, dass das Nutzermodell versucht den Button genau in der Mitte zu treffen. Die Formel für die Berechnung des Breitenfaktors ist äquivalent.

$$\frac{h_B}{2 * z_{FR+}} = f_H \quad (3)$$

Um zu berechnen ob ein anderer Button oder der Hintergrund der Nutzerschnittstelle getroffen werden, werden die ermittelten x und y-Werte mit den Faktoren multipliziert. Der aus der Berechnung resultierende Punkt $P_{x,y}$ wird dann mit den Positionen der anderen Interaktionsobjekte verglichen. Befindet sich der $P_{x,y}$ innerhalb eines anderen Interaktionsobjektes, wird die Interaktion auf dieses Interaktionsobjekt angewendet, sonst gilt die Interaktion auf den Hintergrund der App angewendet.

Die Aufgabe des EMs kann wie folgt zusammengefasst werden. Das EM dient dazu Interaktionsfehler, die sogenannten Slips, zu simulieren. Für die Berechnung ob ein Slip auftritt werden Wahrscheinlichkeiten genutzt, die wiederum von der Größe des Buttons und dem Abstand zu anderen Button abhängen. In dem nächsten Schritt muss nun die Interaktion vom Systemmodell verarbeitet werden.

4.2.6 Modul zur Verarbeitung der Nutzereingabe

Da das mentale Modell des Nutzermodells vom modellierten Systemmodell abweichen kann, muss die Eingabe des Nutzermodells vom System interpretiert werden. Hierfür wurde ein neues Modul zur Verarbeitung der Nutzereingabe (engl. System-Understanding-Modul | SUM) entwickelt.

Das in dieser Arbeit entwickelte SUM ist in der Lage, den Fehler (Ergebnis: *Unbekannte Interaktion*) für auf kapazitive Touchscreens angewendete Interaktionen zu berechnen. Um die empirisch ermittelten Wahrscheinlichkeiten für die Berechnung ob ein Fehler auftreten wird oder nicht auszuwählen, muss vom SUM bestimmt werden, ob die auf den Touchscreen aufgebrachte Interaktion *kurz* oder *lang* und *mit* oder *ohne Druck* ausgeführt wird.

Um zu entscheiden, ob eine Interaktion als lang bewertet wird, wird die Zeit bestimmt, die die Berührung auf dem Touchscreen dauert. Für das SUM gilt, dauert die Berührung länger als 500 ms so wird die Interaktion als lang bewertet. Der Wert von 500 ms leitet sich, wie bereits erläutert, von der Android-Funktion `getLongPressTimeout()` ab [6]. Um zu entscheiden ob eine Interaktion mit Druck oder ohne Druck ausgeführt wird, müssen der aufgebrachte Druck bestimmt und ein Schwellwert definiert werden. In der Literatur existiert bisher

Interaktionstyp	Simulierte Dauer	Simulierter Druck
Short tap	100 ms	0,1
Long tap	500 ms	0,1
Short press	100 ms	0,9
Long press	500 ms	0,9

Tabelle 28: Definierte Werte für die Dauer und den Druck

kein einheitlicher Wert der definiert ab wann aus einem Tap ein Press wird. Wilson und Kollegen [252] nutzten ein Nokia N810, um in einem Experiment verschiedene druckabhängige Interaktionsgesten zu testen. Das Nokia N810 liefert abhängig vom Druck auf dem Bildschirm einen Wert zwischen 0 (0 N (Newton)) und 1 (12 N). Für das SUM bedeutet dies, dass eine Interaktion als mit Druck definiert wird, wenn der Sensorwert größer als bzw. gleich 0,5 ist. Die Tabelle 28 zeigt die in der Werkbank hinterlegten Werte für die vier annotierten Interaktionstypen *Short tap*, *Long tap*, *Short press* und *Long press*. Die Werte für eine kurze Interaktion werden mit 100 ms definiert, dies ist nach [4] die Zeit, die benötigt wird, um eine Taste zu drücken. Für eine lange Interaktion werden die 500 ms verwendet. Für die Druck-Sensorwerte ist eine empirisch hergeleitete Definition nicht möglich, daher wird der Wert für einen Tap mit 0,1 definiert und für einen Press mit 0,9.

Nachdem die Interaktionstypen für die Dimensionen Zeit und Druck definiert sind, kann die jeweilige Wahrscheinlichkeit abgeleitet werden, dass eine Interaktion vom SUM auf eine vom System ausführbare Interaktion transformiert werden kann. Da für den dargestellten empirischen Versuch nur *Short tap*-Interaktionen vom System verstanden worden sind, müssen die alternativen Interaktionstypen mit Wahrscheinlichkeiten belegt werden, dass diese als *Short tap* vom System verarbeitet werden. Die Tabelle 29 stellt die Wahrscheinlichkeiten dar, dass eine Interaktion vom System nicht in eine *Short tap*-Eingabe transformiert werden kann.

Für den Interaktionstyp *Short tap* wurde davon ausgegangen, dass dieser immer als *Short tap* vom System verstanden wird. Beim Interaktionstyp *Long tap* zeigten sich in den empirischen Daten keine Fehler, weshalb auch hier eine Fehlerwahrscheinlichkeit von 0 definiert wurde. Der Interaktionstyp *Short press* führte im Durchschnitt bei jeder 5. Interaktion zu keiner Systemreaktion, weshalb dort eine Fehlerwahrscheinlichkeit von 0,208 definiert ist. Für *Long press* ist die Fehlerwahrscheinlichkeit höher. Bei diesem Interaktionstyp wird knapp jede zweite Interaktion, für das Systemmodell aus dem empi-

Simulierte Interaktionsdauer	Simulierter Druck	p_{Fehler}
100 ms	0,1	0,0
500 ms	0,1	0,0
100 ms	0,9	0,208
500 ms	0,9	0,479

Tabelle 29: Wahrscheinlichkeiten, dass eine Interaktion nicht erkannt wird

rischen Versuch, nicht als *Short tap* verstanden und führt somit zum Ergebnis *Unbekannte Interaktion*.

Die Berechnung, ob eine Interaktion in einen *Short tap* transformiert werden kann oder nicht erfolgt, wie in den anderen Modulen auch, durch den Aufruf eines Zufallsalgorithmus der anhand der gewichteten Wahrscheinlichkeiten eine Entscheidung trifft.

Zusammenfassung zum SUM: Das SUM dient dazu die vom Nutzermodell ausgeführte Interaktion zu interpretieren, bzw. die ausgeführte Interaktion in eine ausführbare Interaktion zu transformieren. Für die App aus dem empirischen Versuch kann das SUM zwei verschiedene Ergebnisse (*Short tap* und *Unbekannte Interaktion*) liefern. Die Wahrscheinlichkeit, dass eine Interaktion nicht in einen *Short tap* transformiert werden kann, hängt von der Dauer und dem aufgewendeten Druck ab.

4.2.7 Modul zum Export der Simulationsergebnisse

Neben den bereits beschriebenen Modulen, die die Simulation direkt beeinflussen, wird auch ein Modul benötigt, welches die Ergebnisse der Simulation ausgeben kann, so dass diese von einem Statistikprogramm verarbeitet werden kann. Hierfür wurde ein neues erweitertes Excel-Export-Modul (EEM) entwickelt.

Das EEM exportiert die Simulationsdaten in ein Excel-Files, welches über zwei Arbeitsblätter verfügt. Arbeitsblatt 1, mit *Raw* beschriftet, enthält die Rohdaten der Interaktion und Arbeitsblatt 2, welches mit *UserData* beschriftet ist, enthält per Iteration zusammengefasste Daten. Für eine Beispielausgabe des EEM wird auf den Anhang B verwiesen, da die exportierten Tabellen sehr breit sind. Die durch das neue EEM erzeugten Logfiles ermöglichen eine einfache und schnelle Auswertung der erzeugten Simulationsergebnisse in Excel, können aber auch in anderen Programmen weiter ausgewertet werden.



Abbildung 38: Views aus der verwendeten App

4.3 SIMULATION

Das Ziel der Simulation ist es zu zeigen, dass es durch die Verwendung der neuen Module möglich ist, die Ergebnisse aus dem empirischen Versuch zu reproduzieren. Hierfür wird im nachfolgenden Abschnitt beschrieben, wie das Systemmodell der App aufgebaut ist und welche Bedingungen für die Simulation gelten. Das Ziel dieser Beschreibung ist es die Ergebnisse der Simulation reproduzierbar und nachvollziehbar zu machen. An die Beschreibung des Modellaufbaus schließt sich der Abschnitt zum Vergleich der Ergebnisse an, der die mit der Simulation erzielten Ergebnisse darstellt und diese mit den im empirischen Versuch gefundenen Ergebnissen vergleicht.

4.3.1 Aufbau des Modells und Bedingungen für die Simulation

In diesem Abschnitt werden der Aufbau des Systemmodells und die Bedingungen für die Simulation beschrieben, wobei der Fokus auf der Beschreibung der Bedingungen liegt, da der schematische Aufbau eines Systemmodells bereits in Abschnitt 2.1 dieser Arbeit beschrieben wurde.

Für die Simulation wurde die im Versuch verwendete App nicht vollständig als Modell umgesetzt. So wurde darauf verzichtet, die Untermenüs der Lampen-, Rollo- und Radio-Steuerung (View *Radio* in Abbildung 38a) vollständig zu modellieren. Vielmehr wurde in den entsprechenden Ansichten nur der Zurück-Button – der Button in der rechten oberen Ecke – modelliert, damit das Nutzermodell auf den Startbildschirm wechseln konnte. Durch diese gewählte Restriktion hat das Nutzermodell nur eine Interaktionsoption auf den entsprechenden Views zur Auswahl und es kann keine *falsche Interaktion* wählen.

Die Begründung der Entscheidung soll an einem Beispiel erklärt werden. Es ist anzunehmen, dass das Nutzermodell die Aufgabe hat den Fernseher leiser zu machen. Wie im Abschnitt 2.2.1 beschrieben, wird das Interaktionsobjekt anhand von Wahrscheinlichkeiten aus-

gewählt. Die Wahrscheinlichkeit ist in der Standardeinstellung nicht absolut – es existiert immer die Wahrscheinlichkeit, dass ein anderes Interaktionsobjekt vom Nutzermodell gewählt wird. Der Ablauf in der Simulation und das resultierende Ergebnis pro Iteration sind somit nicht deterministisch. Wegen des fehlenden Determinismus kann es zu fehlerbehafteten Entscheidungen bei der Auswahl des Interaktionsobjektes kommen, z. B. entscheidet sich das Nutzermodell für den Radio-Button, statt für den Fernseher-Button in Abbildung 38d. Da das Nutzermodell, in der zugrundeliegenden Implementierung, nicht in der Lage ist den Kontext der Benutzerschnittstelle zu interpretieren, kann das Benutzermodell im genannten Beispiel nicht erkennen, dass es sich im vorherigen Interaktionsschritt für das falsche Interaktionsobjekt entschieden hat. Da diese Erkenntnis fehlt, würde das Nutzermodell den Button *Leiser* in Abbildung 38a als relevant einstufen und mit hoher Wahrscheinlichkeit ausführen, statt wie zu erwarten einfach nur den View zu schließen und auf den Startbildschirm zurückzukehren. Um dieses Interaktionsverhalten zu vermeiden, wurden die entsprechenden Views nicht komplett modelliert.

Für den View *Liste der Fernsehsender* (siehe Abbildung 38b), wurde nicht modelliert, welcher Kanal der nächste bzw. der vorherige Kanal ist. Ein Nutzermodell würde deshalb mit hoher Wahrscheinlichkeit den Button *Zurück* wählen, um auf den vorhergehenden View *Fernseher* (Abbildung 38c) zu wechseln. Die Begründung für diese Einschränkung ist, wie schon im vorherigen Beispiel, das Fehlen der Fähigkeit den Kontext zu erkennen.

Bei allen Zurück-Buttons wurde die Größe als *klein* (*small*) modelliert. Die Listenelemente in den einzelnen View wurden als *normal groß* definiert und die vier Hauptbuttons auf dem Startbildschirm der App (Abbildung 38d) als *groß* gekennzeichnet. Die vier Buttons der oberen Leiste auf dem Startbildschirm wurden, ebenso wie die drei Buttons für Video, MP3 und Playlist, als *klein* definiert.

Für keinen Button wurden Kontrastwerte modelliert, so dass alle Buttons einen mittleren Kontrast haben. Weiterhin wurden keine Input- oder Output-Informationen modelliert (vgl. Abschnitt 2.1.3.1). Für die Transitionen zwischen den einzelnen Systemzuständen wurde als Feedback immer ein *explizites Feedback* mit einem *mittleren Level* angenommen (die Standardeinstellungen für Feedback). Die einzigen Ausnahmen sind die *Leiser* und *Lauter* Buttons auf dem View *Fernseher*. Hier wurde das Feedback als *implizit* modelliert und der Level als *sehr gering* (*very low*) definiert. Mit dieser Restriktion wird dem Umstand Rechnung getragen, dass viele Versuchspersonen während des empirischen Versuchs die Unterschiede in der Lautstärke nicht wahrgenommen haben. Es ist anzumerken, dass die Buttons für *lauter* und *leiser* auf dem Startbildschirm mit *explizitem Feedback* und dem *mittleren Level* modelliert wurden. Dies wurde deshalb so definiert, da die Bedienung dieser Buttons zu einer Bildschirmanzeige führte,

Nr.	Teilaufgabe	Wissen: Gerät	Wissen: Aufgabe
1	Fernseher einschalten	Fernseher	einschalten
2	Umschalten auf nächsten Kanal	Fernseher	nächster Kanal
3	Lautstärke verringern	Fernseher	leiser
4	Umschalten auf vorherigen Kanal	Fernseher	vorheriger Kanal
5	Lautstärke erhöhen	Fernseher	lauter
6	Fernseher ausschalten	Fernseher	ausschalten

Tabelle 30: Modelliertes Aufgabenwissen in der Simulation

auf welchem die Versuchsperson auswählen musste ob die Lautstärke des Radios oder des Fernsehers verändert werden soll. Hatte die Versuchsperson das entsprechende Gerät ausgewählt schloss sich der Auswahlbildschirm, so dass das Feedback als explizit und deutlich einzustufen ist.

4.3.2 Aufgabenwissen der Nutzermodelle

Neben den modellierten Simulationsbedingungen spielt auch das modellierte Nutzerwissen eine Rolle, wenn die Ergebnisse nachvollziehbar und wiederholbar sein sollen.

Für die Simulation zur Bedienung des Fernsehers wurde die Aufgabenstellung in 6 Teilaufgaben untergliedert, welche nacheinander von den Nutzermodellen ausgeführt werden. Die erste Teilaufgabe dient dazu den Fernseher einzuschalten. Der Ausgangszustand für diese Teilaufgabe ist wie folgt definiert; auf dem simulierten Touchscreen wird der Startbildschirm (Abbildung 38d) angezeigt und der simulierte Fernseher ist ausgeschaltet. Zur Lösung der Aufgabe hat ein Nutzermodell das Wissen: *Gerät = Fernseher; Aufgabe = einschalten*. Die Teilaufgabe ist erfolgreich gelöst, wenn der simulierte Fernseher eingeschaltet wurde.

In Tabelle 30 ist das Aufgabenwissen für alle 6 Teilaufgaben aufgeführt. Jede Teilaufgabe begann in dem Zustand, in dem die vorhergehende Aufgabe endete. So war der Startbildschirm für die Teilaufgabe 2 der Startzustand, weil nach dem Einschalten des Fernsehers die App automatisch in diesen Zustand wechselte. Für die Teilaufgabe 4 war entweder der View *Fernseher* oder der Startbildschirm der Startzustand, da die Änderung der Lautstärke über den Startbildschirm oder über den View *Fernseher* geschehen konnte und diese beiden Interaktionspfade zu unterschiedlichen Endzuständen führten.

Nachfolgenden werden nun die Resultate der Simulation dargestellt und mit den Ergebnissen aus dem Versuch verglichen.

4.3.3 *Vergleich der empirischen Ergebnisse mit den Simulationsergebnissen*

Das Ziel des Vergleiches ist es zu zeigen, dass es durch die Verwendung der neuen Module möglich ist, die Ergebnisse aus dem empirischen Versuch zu reproduzieren. Nachfolgende werden in den einzelnen Abschnitten die erzeugten Nutzereigenschaften, die simulierte Interaktion, die generierten Interaktionspfade und die unnötigen Interaktionen mit den Ergebnissen aus dem empirischen Versuch verglichen.

4.3.3.1 *Nutzerattribute*

Da die Ausprägung der Nutzerattribute einen entscheidenden Einfluss darauf hat, wie die mentalen Modelle ausgewählt werden, werden in diesem ersten Vergleich die für die Simulation erzeugten Nutzermodelle und deren Attribute mit denen des empirischen Versuchs verglichen. Da die Simulation die Ergebnisse aus dem empirischen Versuch reproduzieren soll, müssen die erzeugten Nutzerattribute den empirischen Nutzerattributen entsprechen.

Die Tabelle 31 stellt die Altersverteilung in den empirischen Daten und aus den Ergebnissen der Simulation dar. Alle Vergleiche der Unterschiede zwischen den Nutzerklassen in den empirischen und den simulierten Daten sind nicht signifikant. Im Vergleich zu der Beschreibung der Versuchspersonen in Unterabschnitt 4.1.3 fällt auf, dass sich die Anzahl der Versuchspersonen in den jüngeren Versuchspersonengruppen jeweils um eine Person verringert hat. Eine jüngere, weibliche Versuchsperson wurde aus der Analyse herausgenommen, da sie die Aufgabenstellung nicht komplett löste und eine jüngere, männliche Versuchsperson wurde nicht in die Analyse eingeschlossen, da nicht alle Interaktionen im Video annotiert werden konnten.

Die Abbildung 39 zeigt die Verteilungen für die Dimensionen der Technikaffinität in den vier Nutzermodellgruppen, welche durch die Simulation erzeugt wurden. Für die Simulation wurden je 500 Iterationen für die Gruppe der älteren Nutzermodelle und 500 Iterationen für die Gruppe der jüngeren Nutzermodelle durchlaufen. Da die Anzahl der Nutzermodelle pro Versuchsgruppe hoch ist, zeigt der Kruskal-Wallis-Rangsummentest für jede Dimension hoch signifikante Unterschiede zwischen den Nutzermodellgruppen.

Die Post-hoc-Vergleiche mit paarweisen Wilcoxon-Tests mit Bonferroni-Holm-Korrektur, zeigten in allen Dimensionen zwischen allen Gruppen hoch signifikante Unterschiede. Ausnahmen waren in der Dimension Kompetenz der Vergleich zwischen älteren Männern und jüngeren Frauen ($p = 0,49$) und in der Dimensionen negative Folgen

Gruppe	Empirischer Versuch			Simulation			<i>p</i>
	n	M	SD	n	M	SD	
Ältere Männer	5	58,2	4,438	185	58,422	3,813	0,850
Ältere Frauen	7	59,286	3,904	315	58,498	4,514	0,829
Jüngere Männer	7	23	2,236	242	22,872	2,204	0,810
Jüngere Frauen	7	25,143	2,795	258	25,271	2,609	0,806

Tabelle 31: Verteilung von Alter und Geschlecht in Simulation und in den empirischen Daten

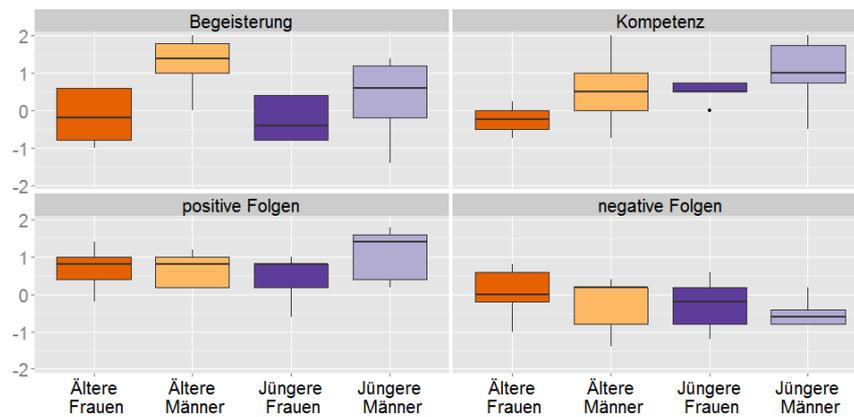


Abbildung 39: Verteilung in den Dimensionen der Technikaffinität für die 4 Nutzermodellklassen. Alle Unterschiede zwischen den Nutzermodellgruppen sind signifikant, bis auf:

- Ältere Männern und jüngere Frauen in der Dimension Kompetenz
- Ältere Männern und jüngere Frauen in der Dimension Negative Einstellung zur Technik
- Ältere Männern und jüngere Männer in der Dimension Negative Einstellung zur Technik

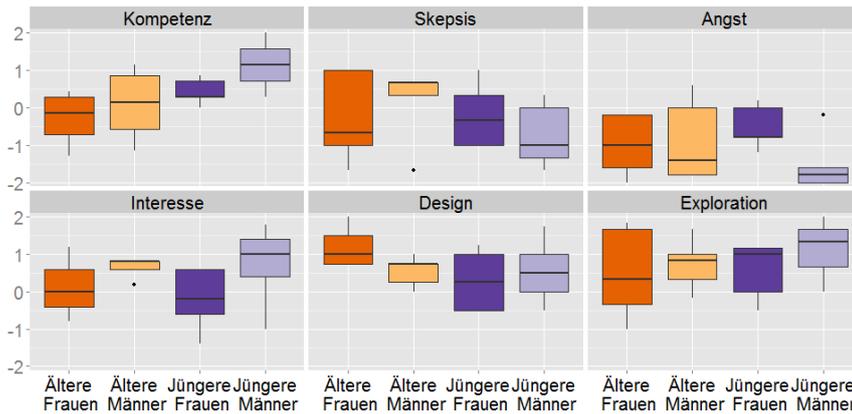


Abbildung 40: Verteilung in den Dimensionen der Technikeinstellung für die 4 Nutzermodellklassen. Alle Vergleiche zwischen den Nutzermodellklassen, bis auf Ausnahmen (siehe Tabelle 32), signifikant.

der Techniknutzung die Vergleiche zwischen älteren Männern und jüngeren Frauen ($p = 0,386$) und älteren Männern und jüngeren Männern ($p = 0,386$).

Der Vergleich der empirischen Daten (siehe Abbildung 27) und den simulierten Daten zeigt für keine Nutzermodellklasse/Versuchspersonengruppe in keiner der vier Dimensionen signifikante Unterschiede. Es kann somit davon ausgegangen werden, dass die durch die Werkbank generierten Attribute für die Technikaffinität denen des empirischen Versuchs entsprechen.

In Abbildung 40 sind die Verteilungen der Nutzermodellattribute für die 4 Nutzermodellklassen und die 6 Dimensionen des ICT-Fragebogens dargestellt. Wie schon für die Werte zur Technikaffinität existieren für alle Dimensionen hochsignifikante Unterschiede zwischen den Nutzermodellklassen. In Tabelle 32 sind die Vergleiche aufgeführt, welche mit dem paarweisen Wilcoxon-Test mit Bonferroni-Holm-Korrektur keine signifikanten Unterschiede aufwiesen.

Der Vergleich der empirischen Daten (siehe Abschnitt 4.1.7.1) und den simulierten Daten zeigte in keiner der sechs Dimensionen signifikante Unterschiede für eine der vier Gruppen.

Da die von der Werkbank erzeugten Nutzerattribute weder für Technikaffinität (TA) noch für Technikeinstellung (ICT) signifikant von den empirischen Daten abweichen, kann ausgeschlossen werden, dass Unterschiede in den Vorhersagen für die Interaktionstypen im nächsten Abschnitt aus den verwendeten Nutzerattributen resultieren.

4.3.3.2 Interaktionen

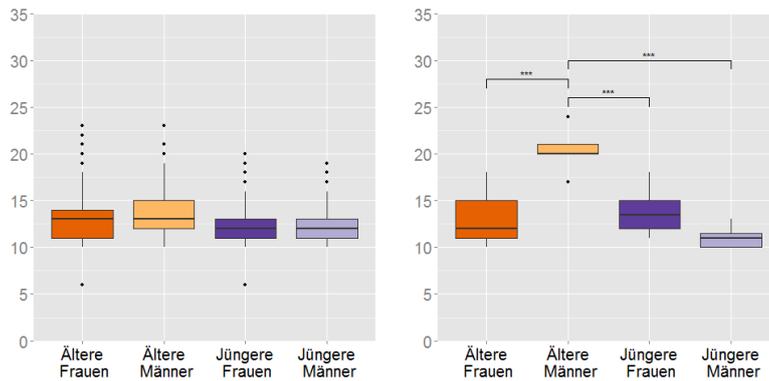
In diesem Unterabschnitt werden die benötigten Interaktionsschritte, die von den Nutzermodellen verwendeten Interaktionstypen und die

Dimension	Nutzerklasse 1	Nutzerklasse 2	<i>p</i>
Exploration	Ältere Frauen	Ältere Männer	0,144
Exploration	Ältere Frauen	Jüngere Frauen	0,626
Exploration	Ältere Männer	Jüngere Frauen	0,626
Design	Ältere Männer	Jüngere Männer	0,4
Interesse	Ältere Frauen	Jüngere Frauen	0,16
Angst	Ältere Männer	Ältere Frauen	0,52
Skepsis	Ältere Frauen	Jüngere Frauen	0,358

Tabelle 32: Keine Unterschiede zwischen den Nutzermodellgruppen in den Dimensionen

Interaktionspfade dargestellt und mit den empirischen Daten verglichen. Da nicht alle Versuchspersonen alle Aufgaben im empirischen Test durchführen konnten, wird für den Vergleich die Unteraufgabe TV ausgewählt. Die nachfolgenden Angaben beziehen sich somit auf Interaktionen, die mit dem Fernseher durchgeführt wurden. Wie bereits in Unterabschnitt 4.1.5 dargestellt waren zum Lösen der Aufgabenstellung mindesten 10 Interaktionsschritte notwendig.

In Abbildung 41 sind die Verteilungen für die benötigten Schritte im TV Task für die 4 verschiedenen Nutzerklassen dargestellt. Für die empirischen Daten (Abbildung 41b) existieren hochsignifikante Unterschiede zwischen den verschiedenen Versuchspersonengruppen ($F_{df}(3,21) = 16,04; p < 0,001$). Der eingesetzte Post-hoc-T-Test mit Bonferroni-Holm-Korrektur zeigt Unterschiede zwischen älteren Männern und älteren Frauen ($p < 0,001$), älteren Männern und jüngeren Frauen ($p < 0,001$) und älteren Männern und jüngeren Männern ($p < 0,001$). Für alle anderen Paarvergleich ist der Test nicht signifikant. Bei den simulierten Daten (Abbildung 41a) zeigt der Kruskal-Wallis-Rangsummentest signifikante Unterschiede ($\chi^2 = 75,58; df = 3; p < 0,001$) zwischen den Nutzermodellgruppen hinsichtlich der benötigten Interaktionsschritte. Der als Post-hoc durchgeführt korrigierte Wilcoxon-Test zeigt signifikante Unterschiede zwischen allen Gruppen, nur der Unterschied zwischen jüngeren Frauen und jüngeren Männern ist nicht signifikant. Beim Vergleich der empirischen Daten mit den simulierten Daten zeigt der Mann-Whitney-U-Test signifikante Unterschiede ($p = 0,003$) zwischen den simulierten und empirischen Daten für die Nutzerklasse der älteren Männer. Im empirischen Versuch haben die älteren Männer deutlich mehr Interaktionen benötigt, um die Aufgaben zu lösen. Die detaillierte Auswertung zeigte, dass die älteren Männer im empirischen Versuch häufiger Interaktionen



(a) Simulation: Bis auf den Vergleich zwischen jüngeren Männern und jüngeren Frauen, sind alle signifikant unterschiedlich

(b) Empirischer Versuch

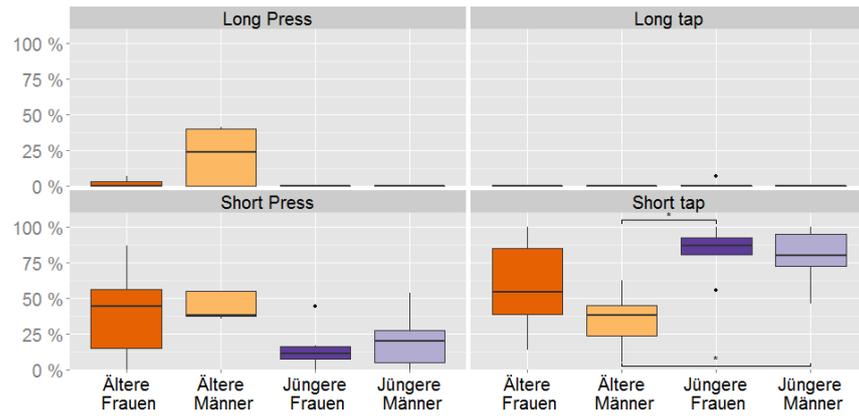
Abbildung 41: Anzahl der benötigten Schritte

mit schlechtem Feedback wiederholten, als dies die simulierten älteren männlichen Nutzermodelle taten.

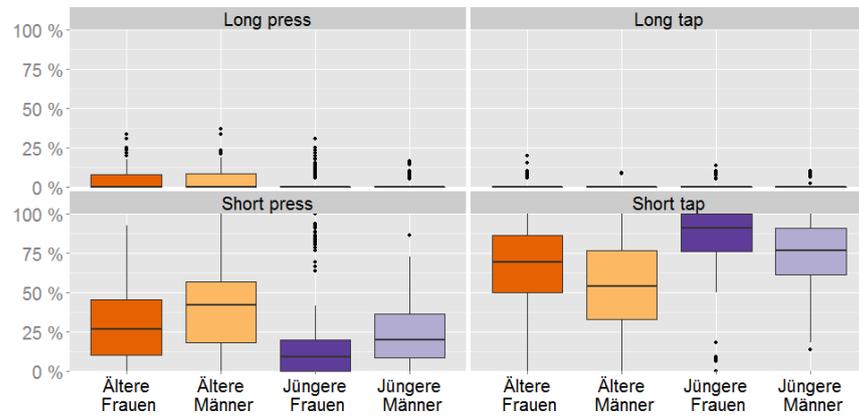
In Abbildung 42a sind die Häufigkeiten für die verschiedenen Interaktionstypen in den vier Versuchspersonengruppen (empirischer Versuch) dargestellt. Der Vergleich zwischen den Versuchspersonengruppen für die einzelnen Interaktionstypen brachte, bis auf den Vergleich für *Long tap*, immer signifikante Unterschiede zwischen den Gruppen. Bei den anschließenden Post-hoc-Tests (paarweiser T-Test bzw. Mann-Whitney-U-Test wenn keine Homoskedastizität vorlag) zeigten sich allerdings nur für *Short tap* signifikante Unterschiede. So kann festgestellt werden, dass jüngere Frauen bzw. jüngere Männer *Short tap* signifikant häufiger nutzen als dies ältere Männer tun.

In Abbildung 42b sind die Häufigkeiten der verwendeten Interaktionstypen für die unterschiedlichen Nutzermodellklassen (Simulation) dargestellt. Der Vergleich der Häufigkeiten für die verschiedenen Interaktionstypen zwischen den Nutzermodellklassen ergab für alle Interaktionstypen, außer für *Long tap*, hochsignifikante Unterschiede zwischen den Nutzermodellklassen. Der Post-hoc-Test (Mann-Whitney-U-Test) zum paarweisen Vergleich zeigte signifikante Unterschiede bei allen Vergleichen. Beim Vergleich zwischen Simulation und empirischem Versuch hinsichtlich der verwendeten Interaktionstypen zeigten sich keine signifikanten Unterschiede in den Häufigkeiten für die Versuchspersonengruppen bzw. Nutzermodellklassen.

Neben der Anzahl der benötigten Interaktionsschritte und den verwendeten Interaktionstypen sind besonders die durch die Simulation erzeugten Interaktionssequenzen für eine Auswertung interessant. Die nachfolgende Tabelle 33 führt die Interaktionssequenzen für die jüngeren männlichen Versuchspersonen und deren Häufigkeiten auf. In der Spalte *Interaktionssequenzen* steht, welche Interaktionen aufeinander



(a) Empirischer Versuch



(b) Simulation

Abbildung 42: Häufigkeit der Verwendung der Interaktionstypen im TV Task

folgen. In der Spalte *Anzahl* ist die absolute Häufigkeit des Auftretens dieser Sequenzen aufgeführt und in der Spalte *Anteil* die prozentuale Verteilung. Die Einträge in der Spalte *Interaktionssequenzen* sind wie folgt zu lesen; *Fernseher/1* bedeutet, dass der Button *Fernseher* auf dem Startbildschirm der App (siehe Abbildung 38d) einmal bedient wurde. Der Eintrag *Einschalten/1* bedeutet, dass der Button *Einschalten* (siehe Abbildung 38c) einmal benutzt wurde. Die Einträge *Nächster Sender*, *Leiser*, *Vorheriger Sender*, *Lauter* und *Ausschalten* stehen äquivalent für die entsprechenden Buttons. Ist die Zahl hinter dem Schrägstrich größer als 1, so wurde die Aktion mehrmals ausgeführt; ein *Leiser/2* bedeutet, dass der Button *Leiser* zweimal hintereinander ausgeführt wurde. In der Tabelle 33 sind unnötige Wiederholungen, aufgrund von mangelhaftem Feedback, mit roter Farbe markiert. Mit Gelb sind Interaktionen markiert, die nicht notwendig oder fehlerbehaftet waren, allerdings auch nicht auf mangelhaftes Feedback zurückzuführen sind. Wenn die Interaktionssequenzen in Tabelle 33 betrachtet werden, so fällt auf, dass die Unterschiede hauptsächlich in der Wiederholung von Interaktionen zur Veränderung der Lautstärke liegen und nur eine Versuchsperson die Reihenfolge der Aufgaben vertauscht.

Die nachfolgende Tabelle 34 enthält die Interaktionssequenzen und deren Häufigkeiten für die simulierten jüngeren männlichen Nutzermodelle. Wie in den empirischen Daten ist die optimale Interaktion, die Lösung der Aufgabe mit 10 Schritten, die am häufigsten vorzufindende Interaktionssequenz. Werden die Interaktionssequenzen in Tabelle 34 miteinander verglichen, so fällt auf, dass sich die häufigsten Sequenzen nur hinsichtlich der Wiederholungen aufgrund von mangelhaftem Feedback (rot markiert *Leiser* und *Lauter*) unterscheiden und das keine Vertauschung der Reihenfolge vorkommt. Dieses Verhalten ist gewollt bzw. ein Algorithmus zur Aufgabenvertauschung ist bisher nicht implementiert.

In Tabelle 35 sind die Interaktionssequenzen für die älteren männlichen Versuchspersonen dargestellt. Wie in der Tabelle zu sehen ist, sind keine zwei Sequenzen identisch. Werden die Interaktionssequenzen der älteren Männer mit denen der jüngeren männlichen Versuchspersonen verglichen, so fällt auf, dass die älteren Männer häufiger Interaktionen wiederholen. Ein Großteil der Wiederholungen ist auf das mangelnde Feedback zurückzuführen (rot markiert). Mit Blau markierte Einträge sind durch Interaktionsfehler entstanden, hier wurde die Eingabe aufgrund eines *Slips* oder *Mistakes* nicht vom System verarbeitet. Die gelb markierten Einträge sind unnötige Wiederholungen bereits erledigter Teilaufgaben.

Vergleicht man die empirischen Daten aus Tabelle 35 mit den simulierten Ergebnissen in Tabelle 36 so fällt auf, dass auch bei den simulierten Ergebnissen keine eindeutig häufigeren Interaktionssequenzen existieren. Die geringen Werte für den Anteil der häufigsten Interaktionssequenzen deuten darauf hin, dass viele verschiedene

Rang	Interaktionssequenz	Häufigkeiten	
		Anzahl	Anteil
1	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Fernseher/1, Lauter/1, Ausschalten/1	2	29 %
2	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Fernseher/1, Leiser/1, Lauter/1, Ausschalten/1	1	14 %
2	Fernseher/1, Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/2, Vorheriger Sender/1, Fernseher/1, Lauter/1, Ausschalten/1	1	14 %
2	Fernseher/1, Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/2, Vorheriger Sender/1, Fernseher/1, Lauter/2, Ausschalten/1	1	14 %
2	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/3, Vorheriger Sender/1, Fernseher/1, Lauter/2, Ausschalten/1	1	14 %
2	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Vorheriger Sender/1, Fernseher/1, Leiser/1, Lauter/1, Ausschalten/1	1	14 %

Tabelle 33: Interaktionssequenzen der jüngeren, männlichen Versuchspersonen (empirischer Versuch)

Rang	Interaktionssequenz	Häufigkeiten	
		Anzahl	Anteil
1	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Fernseher/1, Lauter/1, Ausschalten/1	37	15,3 %
2	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Fernseher/1, Lauter/2, Ausschalten/1	19	7,9 %
3	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/2, Vorheriger Sender/1, Fernseher/1, Lauter/1, Ausschalten/1	16	6,6 %
4	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Fernseher/1, Lauter/3, Ausschalten/1	13	5,4 %
5	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/2, Vorheriger Sender/1, Fernseher/1, Lauter/2, Ausschalten/1	8	3,3 %

Tabelle 34: Interaktionssequenzen der jüngeren, männlichen Nutzermodelle (Simulation)

Rang	Interaktionssequenz	Häufigkeiten	
		Anzahl	Anteil
1	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/8, Lauter/2, Vorheriger Sender/1, Fernseher/1, Lauter/2, Ausschalten/1	1	20 %
1	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/2, Fernseher/1, Leiser/5, Lauter/2, Ausschalten/3, Listenansicht schließen/1, Fernseher/1, Ausschalten/3	1	20 %
1	Fernseher/1, Einschalten/3, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/5, Vorheriger Sender/5, Fernseher/1, Lauter/2	1	20 %
1	Fernseher/1, Einschalten/4, Fernseher/1, Nächster Sender/1, Fernseher/1, Vorheriger Sender/4, Fernseher/1, Lauter/1, Ausschalten/1	1	20 %
1	Fernseher/1, Nächster Sender/2, Einschalten/2, Fernseher/1, Vorheriger Sender/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Lauter/1, Wahl Fernseher/1, Lauter/1, Wahl Fernseher/1, Mute/1, Wahl Fernseher/1, Fernseher/1, Lauter/2, Ausschalten/1	1	20 %

Tabelle 35: Interaktionssequenzen der älteren, männlichen Versuchspersonen (empirischer Versuch)

Rang	Interaktionssequenz	Häufigkeiten	
		Anzahl	Anteil
1	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Fernseher/1, Lauter/1, Ausschalten/1	3	1,62 %
1	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/1, Vorheriger Sender/1, Fernseher/1, Lauter/3, Ausschalten/1	3	1,62 %
1	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/2, Vorheriger Sender/1, Fernseher/1, Lauter/1, Ausschalten/1	3	1,62 %
4	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/1, Fernseher/1, Leiser/4, Zurück/1, Fernseher/1, Vorheriger Sender/1, Fernseher/1, Lauter/2, Ausschalten/2	2	1,08 %
4	Fernseher/1, Einschalten/1, Fernseher/1, Nächster Sender/3, Fernseher/1, Leiser/2, Vorheriger Sender/1, Fernseher/1, Lauter/2, Ausschalten/1	2	1,08 %

Tabelle 36: Interaktionssequenzen der älteren, männlichen Nutzermodelle (Simulation)

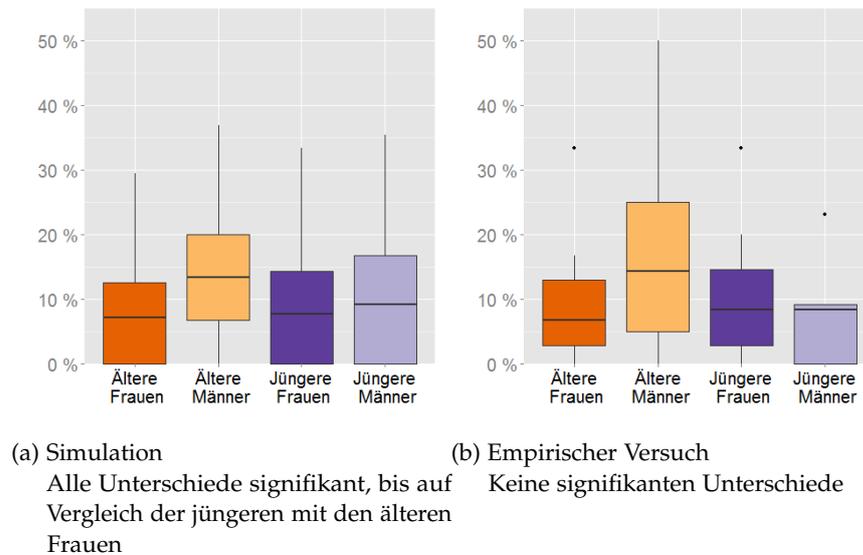


Abbildung 43: Anteil von unnötigen Interaktionen

Interaktionssequenzen durch die Simulation erzeugt wurden. Diese Heterogenität spiegelt sich auch in der Anzahl der unterschiedlichen Interaktionssequenzen wider. Für die 185 älteren männlichen Nutzermodelle wurden insgesamt 172 verschiedene Interaktionssequenzen erzeugt, wohingegen für die 242 jüngeren männlichen Nutzermodelle nur 112 verschiedene Interaktionssequenzen simuliert wurden. In Tabelle 36 sind die Interaktionsfehler ebenfalls farblich markiert. Die Kodierung der Markierungen entspricht der aus Tabelle 35.

4.3.3.3 Unnötige Interaktionen

Wie in Abschnitt 4.1.7 erläutert führt mangelhaftes Feedback zu Wiederholungen von Interaktionen.

In Abbildung 43 sind der Anteil der unnötigen Interaktionen im TV Task im empirischen Versuch und in der Simulation dargestellt. Für die Simulation (Abbildung 43a) existieren, aufgrund der hohen Fallzahl, signifikante Unterschiede zwischen den Nutzermodellgruppen ($\chi^2 = 65,773$; $df = 3$; $p < 0,001$). Die als Post-hoc-Test durchgeführten Paarvergleiche sind alle hochsignifikant, außer für den Vergleich der älteren weiblichen Nutzermodelle mit den jüngeren weiblichen Nutzermodellen. Für den empirischen Versuch (Abbildung 43b) konnten keine signifikanten Unterschiede zwischen den Versuchspersonengruppen gefunden werden. Der Vergleich der Ergebnisse aus der Simulation mit den empirischen Daten ist für keine Versuchspersonengruppe bzw. Nutzermodellklasse signifikant. Die Simulation ist somit in der Lage, die Anteile der unnötigen und notwendigen Interaktionen zu reproduzieren.

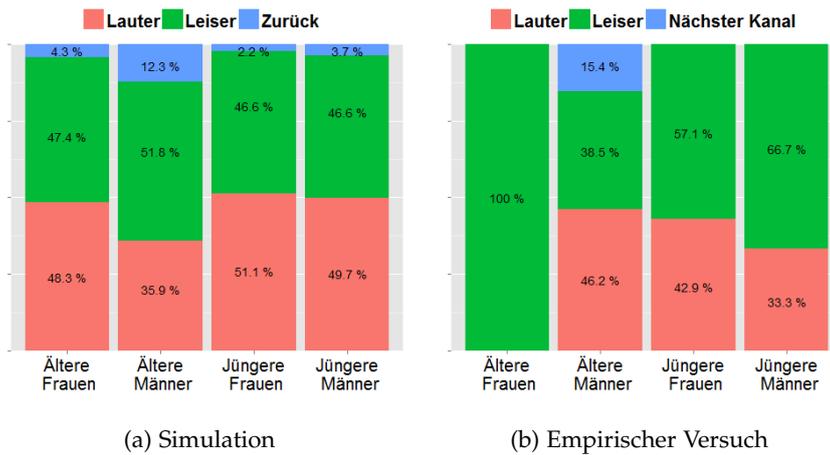


Abbildung 44: Verteilung der Häufigkeiten für unnötige Interaktionen

In Abbildung 44 sind die Anteile der unnötigen Interaktionen in der Simulation (Abbildung 44a) wie auch in den empirischen Daten (Abbildung 44b) dargestellt. Wie den Abbildungen zu entnehmen ist, resultieren für alle Nutzermodellklassen und für alle Versuchspersonengruppen, die unnötigen Interaktionen aus mangelhaftem Feedback (*Lauter* und *Leiser*). In den empirischen Daten hat nur eine Versuchspersonengruppe (ältere Männer) unnötige Interaktionen ausgeführt welche nicht auf mangelhaftes Feedback zurückzuführen ist (*Nächster Kanal*). Für die Simulation zeigt sich ein ähnliches Bild, auch hier überwiegen Wiederholungen aufgrund von nicht wahrgenommenen Feedbacks (*Lauter* und *Leiser*). Unnötige *Zurück*-Interaktionen sind in Interaktionspfaden enthalten, wo das Nutzermodell z. B. den Fernseher ausschalten soll, stattdessen allerdings erst den View Fernseher schließt, um ihn im Anschluss erneut zu öffnen und den Fernseher auszuschalten.

4.4 FAZIT

Die in diesem Kapitel vorgestellte Implementierung kann Fehler aufgrund von falschen mentalen Modellen, unnötigen Wiederholungen und unbeabsichtigten Bedienungen simulieren. Hierfür nutzt die Implementierung verschiedene Algorithmen zur Vorhersage des benutzten mentalen Modells, der Ausführung der Interaktion und dem Systemverstehen. Wie in Abschnitt 4.3.3 gezeigt unterscheiden sich die durch die Simulation erzeugten Ergebnisse nicht signifikant von den Ergebnissen aus dem Versuch. Die Implementierung ist in der Lage, die im empirischen Versuch gefundenen Ergebnisse zu reproduzieren.

4.4.1 *Positive Aspekte*

Die Implementierung ist in der Lage die empirisch beobachteten Interaktionen bzw. Ergebnisse zu reproduzieren, ohne signifikant von diesen abzuweichen. So ist es, wie im empirischen Versuch, die Gruppe der älteren Männer, die im Durchschnitt die meisten Interaktionen benötigt, um die Aufgaben im TV Task zu lösen. Werden die Häufigkeiten der verwendeten Interaktionstypen betrachtet, so ist auch hier eine hohe Übereinstimmung von simulierten und empirischen Ergebnissen zu erkennen. Da die älteren männlichen Versuchspersonen, wie auch die älteren männlichen Nutzermodelle, überdurchschnittlich häufig andere Interaktionen als *Short tap* genutzt haben, hat diese Versuchspersonengruppe/Nutzermodellklasse einen deutlich höheren Anteil an nicht erfolgreichen Interaktionen als dies die anderen Versuchspersonengruppen/Nutzermodellklassen haben.

Hinsichtlich der unnötigen Interaktionen ist festzustellen, dass die Integration des Moduls zur Wahrnehmungen des Systemfeedbacks einen entscheidenden Einfluss darauf hat, dass Interaktionen wiederholt werden und dass deshalb mehr fehlerhafte Interaktionen in den Interaktionssequenzen zu finden sind. Durch die unnötige Wiederholung von Teilaufgaben aufgrund von nicht wahrgenommenen Feedback erscheint das simulierte Interaktionsverhalten realistischer.

Ein weiterer sehr positiver Aspekt ist die Heterogenität der erzeugten Interaktionssequenzen. Speziell die höhere Heterogenität in den Sequenzen für die älteren Nutzermodelle scheint im logisch, da aus der Literatur (siehe [67]) bekannt ist, dass ältere Menschen heterogener in ihren Nutzerattribute sind und hieraus heterogene Interaktionspfade resultieren.

4.4.2 *Negative Aspekte*

Neben den positiven Aspekten der Simulation existieren auch Faktoren, die von der Simulation nicht oder nur schlecht abgedeckt werden.

Ein erster Faktor betrifft das Fehlen von Aufgabenvertauschung bzw. Vertauschungen der Reihenfolge von Teilaufgaben. Die Vertauschung der Reihenfolge von Teilaufgaben ist, wie im empirischen Versuch gesehen, relevant und kann zu interessanten Problemen während der Interaktion führen. Allerdings ist die Umsetzung in der Simulation nicht trivial und die nötige Modellierung der Aufgabenstellungen sehr schwierig. Zum einen müssen Abhängigkeiten unter den Aufgaben definiert werden – z. B. muss der Fernseher erst eingeschaltet werden, bevor man die Lautstärke verändern kann – und zum anderen muss vom Modellierer festgelegt werden, wann Teilaufgaben vertauscht oder wiederholt werden können/müssen. Für beide Herausforderungen muss während der Modellierung beachtet werden, dass keine Informationen fehlen und keine Informationen zu viel modelliert

werden, da sonst die Ergebnisse der Simulation verzerren werden können.

Ein weiterer interessanter Aspekt betrifft das Vergessen von Teilaufgaben. Die bisherige Implementierung ist nicht in der Lage die Ausführung von Teilaufgaben zu vergessen. Die Probleme, die aus dem Vergessen einer Teilaufgabe entstehen sind ähnlich denen der Aufgabenvertauschung. Wieder müssen Abhängigkeiten unter den Teilaufgaben definiert werden und muss bei der Modellierung darauf geachtet werden, dass keine Informationen modelliert werden, die ein Vergessen verhindern bzw. die zu unrealistischen Simulationen führen. Es müssen weiterhin vom Modellierer Annahmen getroffen werden, welche Aufgaben vergessen werden können und welche auf jeden Fall ausgeführt werden.

VALIDIERUNGSVERSUCH

In diesem Kapitel wird ein unabhängiges Experiment zur Validierung der Erweiterung der MeMo-Werkbank beschrieben. Zu Beginn des Kapitels werden das Versuchsdesign und die Zielstellung beschrieben. Im Abschnitt 5.6 werden die Ergebnisse aus dem Versuch mit den Vorhersagen der MeMo-Werkbank und des Bayes'schen Netzes verglichen.

5.1 ZIELSTELLUNG

Wie in Kapitel 4 dargestellt, zeigen ältere Versuchspersonen bei der Steuerung eines Fernsehers mittels Smartphones ein Interaktionsverhalten, welches zu mehr Interaktionsfehlern führt. In diesem Versuch wird untersucht, wie Versuchspersonen aus verschiedenen Altersgruppen und mit unterschiedlichen Vorerfahrungen im Umgang mit Technik, einen Fernseher mit einem Tablet steuern und wie gut sich die empirischen Ergebnisse mit den in dieser Arbeit entwickelten Modellen vorhersagen lassen.

Das Hauptziel des Validierungsversuchs ist es, die Erweiterung der Werkbank aus Kapitel 4 durch ein unabhängiges Experiment überprüfen zu können. Konkret soll gezeigt werden, dass mit der Erweiterung der Werkbank zur Simulation von mentalen Modellen die Interaktionsparameter korrekt simuliert werden können.

Im Zusammenhang mit dem Hauptziel werden zwei Hypothesen durch das Experiment untersucht. Die erste Hypothese beschreibt die Annahme, dass das in Kapitel 4 vorgestellte Modell zur Vorhersage von Interaktionsparametern auch für diesen Versuch gültig ist und dass es die Interaktionsparameter reproduzieren kann. Die zweite Hypothese besagt, dass die Kombination von Bayes'schen Netz und dem Modell für die Vorhersage von Interaktionsparametern, neue und nachvollziehbare Interaktionspfade erzeugt.

1. Hypothese

2. Hypothese

5.2 AUFBAU UND GERÄTE

Der in diesem Kapitel beschriebene Versuch wurde in einer Laborumgebung durchgeführt. Für das Experiment wurden die Probanden aufgefordert sich an einen Tisch zu setzen. Vor ihnen in etwa 2 m Entfernung war ein 40 Zoll großer Fernseher an einem Gestell aufgehängt. Die Probanden wurden während der Interaktion von zwei verschiedenen Kameras aufgenommen. Die erste Kamera filmte den Probanden von links und zeichnete zusätzlich das Fernsehbild auf.



Abbildung 45: Startbildschirm von XBMC auf dem Fernseher.

Die an der Decke befestigte zweite Kamera filmte die Probanden von oben und ermöglichte die Aufzeichnung des Bildschirms des Tablets, welches zur Bedienung des Fernsehers verwendet wurde.

Als Tablet wurde ein Nexus 7 der ersten Generation verwendet. Der berührungsempfindliche Bildschirm des Tablets hat eine Auflösung von 1280 x 800 Pixel bei einer Bildschirmdiagonale von 7 Zoll. Auf dem Nexus 7 ist eine modifizierte Version von Cyanogenmod, einer Variante des Android-Betriebssystems, installiert. In dieser modifizierten Variante können der Bildschirminhalt als Video und die Nutzereingaben als Protokoll, unabhängig von der verwendeten App, aufgezeichnet werden. Als Nutzereingaben werden die Anzahl und die Position der Finger auf dem Bildschirm gespeichert. Zur Erfassung der Werte wird das Betriebssystem-eigene Tool (*dumpsys*) verwendet. Das Tool *Dumpsys* liefert die X und Y-Werte der Finger auf dem Touchscreen und einen Wert für den aufgebrachten Druck [5]. Die X und Y-Werte leiten sich aus der Position der Finger und der Auflösung des Displays ab. Der aufgebrachte Druck wird, wenn vorhanden, aus einem Drucksensor ausgelesen oder auf Basis der Fingerauflagefläche berechnet. Bei der Berechnung des Drucks auf Basis der Auflagefläche ist zu beachten, dass unterschiedliche Personen und unterschiedliche Haltepositionen zu Fehlern in der Berechnung führen können. Neben den Eingaben auf dem Touchscreen, werden auch die seitlichen Buttons zur Änderung der Lautstärke vom System protokolliert. Zusätzlich wird die Orientierung des Tablets im Raum anhand der X, Y, und Z-Koordinaten, inklusive eventueller Winkeländerungen, erfasst.

Auf dem Fernseher wurde die in Abbildung 45 dargestellte Oberfläche von XBMC (mittlerweile in Kodi ⁷ umbenannt) in einer Auflösung

⁷ <http://kodi.tv/>

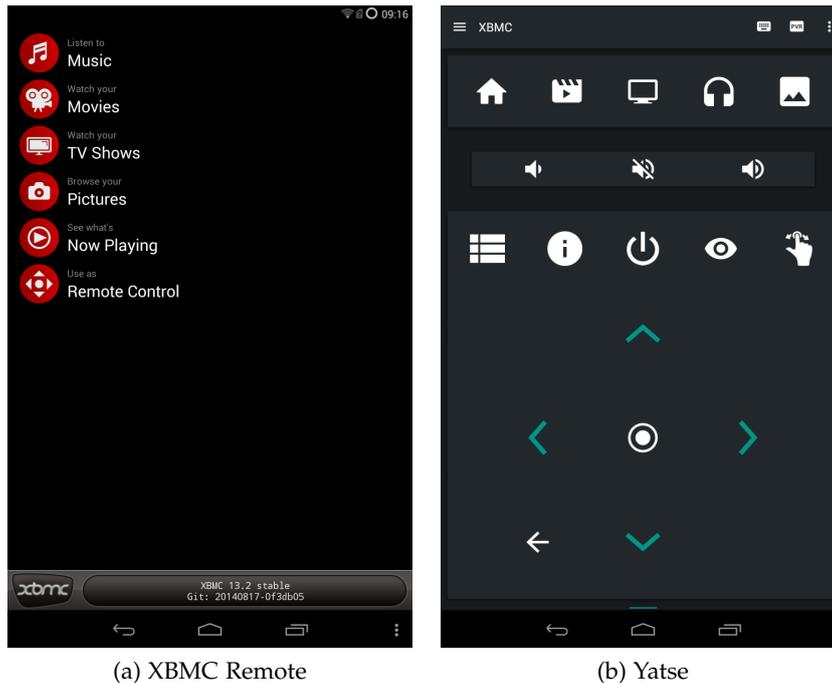


Abbildung 46: Startbildschirme der Android Apps zur Steuerung von XBMC

von 1920 x 1080 Pixel angezeigt. XBMC ist ein offenes und frei verfügbares Mediacenter, das auf verschiedenen Plattformen installiert werden kann. Mit dem Mediacenter können Bilder, Filme, Serien, Musik, und Live-TV abspielen werden. Da XBMC nicht auf einem Fernseher installiert werden kann, wurde es auf einem Laptop eingerichtet. Der Laptop wurde in der Nähe des Fernsehers versteckt und via HDMI mit diesem verbunden. Für den Versuch wurde die Oberfläche in deutscher Sprache verwendet. Die Bedienung des Mediacenters erfolgte über zwei verschiedene Apps, die auf dem Tablet installiert waren.

Die erste App ist die *XBMC Remote*, welche in Abbildung 46a dargestellt ist. Wie in der Abbildung zu erkennen, liegt die App nur in englischer Sprache vor. Auf dem ersten Bildschirm kann der Nutzer per Schnellzugriff auf die Untermenüs für Musik (*Music*), Filme (*Movies*), TV-Serien (*TV Shows*) und Bilder (*Pictures*) zugreifen. Zusätzlich kann über den Menüpunkt *Now Playing* eine Ansicht angezeigt werden, die die aktuelle Wiedergabe im Mediacenter darstellt. Der Menüpunkt *Remote Control* bietet eine Ansicht mit der per Steuerkreuz durch die Oberfläche von XBMC navigiert werden kann.

Die zweite App, die für die Fernsteuerung des Fernsehers verwendet wurde, ist die App *Yatse*, deren Startbildschirm in Abbildung 46b dargestellt ist. Die Yatse App bietet über die in der oberen Leiste dargestellten Piktogramme einen Schnellzugriff auf Filme (🎬), TV-Serien (📺), Musik (🎧) und Bilder (🖼️). Das Haus (🏠) auf der linken Seite bringt den Nutzer auf dem TV Bildschirm zum Startbild der

XBMC-Oberfläche, welcher in Abbildung 45 dargestellt ist. Unter der Leiste für den Schnellzugriff befinden sich die Bedienelemente für die Änderung der Lautstärke. Die dritte Leiste bietet weitere Buttons zur Steuerung des Mediacenters. Der erste Button auf der linken Seite () öffnet in XBMC das Kontextmenü. Der Inhalt des Kontextmenüs hängt davon ab, was aktuell auf der Oberfläche des Mediacenters angezeigt wird. Der Button neben dem Kontextmenü () führt zur Anzeige von Informationen über den aktuellen Status des Mediacenters. In der Mitte der Leiste befindet sich der Button () um das Mediacenter zu beenden bzw. um den PC auszuschalten. Ist der PC ausgeschaltet und kann per Wake-on-Lan geweckt werden, so kann über diese Taste der PC eingeschaltet werden. Der Button mit dem Auge () kann dazu verwendet werden, das Seitenverhältnis (z. B. 16:9 oder 4:3) des Mediacenters zu verändern. Der letzte Button () schaltet zwischen dem Steuerkreuz und einer 2D-Gestensteuerung um. Bei der 2D-Gestensteuerung wird durch Wischen in einem definierten Bereich auf dem Tablet durch die Oberfläche von XBMC navigiert. Wird das Steuerkreuz verwendet, so kann in der Oberfläche des Mediacenters nach *links*, *rechts*, *runter* und *hoch* navigiert werden.

5.3 VERSUCHSPERSONEN

Am Versuch nahmen insgesamt 12 jüngere und 12 ältere Versuchspersonen teil. Die jüngeren Versuchspersonen (18 - 30 Jahre; Median = 26,5 Jahre; Mittelwert = 25,42; SD = 3,85; 6 Frauen) wurden über die Probandendatenbank des Zentrums für Mensch-Maschine-Systeme rekrutiert. Die älteren Teilnehmer (68 - 80 Jahre; Median = 72 Jahre; Mittelwert = 72,58; SD = 3,5; 11 Frauen) wurden, teilweise, ebenfalls über die Probandendatenbank geworben. Da über das Portal nicht genügend ältere Versuchspersonen erreicht werden konnten, wurden verschiedene Seniorenbegegnungsstätten kontaktiert. In den Begegnungsstätten wurden Aushänge und Flyer verteilt, allerdings konnten hierdurch nur wenige zusätzliche Probanden gewonnen werden. Der Großteil der fehlenden älteren Teilnehmer wurde durch die älteren Probanden geworben, die bereits an der Studie teilgenommen hatten.

Die Texte für die E-Mails, Aushänge und Flyer wurden für die älteren und jüngeren Teilnehmer leicht unterschiedlich formuliert. Bei den jüngeren Versuchsteilnehmern wurde geschrieben, dass zwei verschiedene Apps zur Steuerung eines Fernsehers getestet werden sollen. Bei den älteren Versuchspersonen wurden von zwei verschiedenen Fernbedienungen gesprochen. In allen Ausschreibungen wurde darauf hingewiesen, dass keinerlei technisches Wissen nötig ist, um am Versuch teilzunehmen.

Aufgrund der zu erwartenden Herausforderungen eines Usability-Tests unter Beteiligung von älteren Versuchspersonen (siehe Abschnitt 1.3.4), wurden die veranschlagten Durchführungszeiten pro Versuchs-

durchgang und Aufwandsentschädigungen für beide Gruppen unterschiedlich geplant. Bei der jüngeren Gruppe wurde davon ausgegangen, dass ein Versuchsdurchgang maximal 60 Minuten dauert. Aus diesem Grund wurde die Aufwandsentschädigung pro jüngeren Teilnehmer auf 10,- Euro festgelegt. Für die älteren Versuchspersonen wurde eine Durchführungszeit von 90 Minuten eingeplant, weshalb die Vergütung mit 15,- Euro ausgeschrieben wurde.

Von den 24 teilnehmenden Versuchspersonen besaßen zum Zeitpunkt der Teilnahme 21 Personen ein eigenes Mobiltelefon. Die 3 Probanden, die kein Mobiltelefon besaßen, gehören zur Gruppe der älteren Versuchsteilnehmer. Die im Versuch gefundene Verteilung über den Besitz von Mobiltelefonen, findet sich auch in der Literatur. So verzichten, laut einer Studie des Bundesverbands Informationswirtschaft, Telekommunikation und neue Medien e. V. (BITKOM e. V.) aus dem Jahr 2013, 20 % der Personen über 65 Jahren auf ein Mobiltelefon, bei den unter 30-Jährigen verzichtet nur jeder Hundertste [246].

Bei den jüngeren Probanden gaben 100 % an, dass sie mehrmals täglich ein Gerät verwenden, welches über einen berührungsempfindlichen Bildschirm gesteuert wird. Der Grund hierfür ist, dass von den jüngeren Versuchspersonen ausnahmslos Mobiltelefone verwendet werden, die per Touch bedient werden. In der Gruppe der älteren Personen gab nur eine Person (8,3 %) an, ein Gerät mit Touchscreen mehrmals täglich zu nutzen. Eine weitere ältere Person gab an, einmal täglich ein Gerät über einen Touchscreen zu steuern. Beide Personen, die mindestens einmal täglich einen berührungsempfindlichen Bildschirm bedienen, besitzen ein Mobiltelefon mit einem Touchscreen. 3 Personen gaben an, 1 bis 3 Mal pro Woche mit touch-basierten Eingabegeräten zu interagieren und 6 Personen (50 %) sagten aus, dass sie ein Gerät mit Touch-Eingabe seltener als einmal pro Monat verwenden. Eine ältere Person machte keine Angaben zur Verwendungshäufigkeit.

In der Tabelle 37 ist aufgeschlüsselt mit welchen touch-basierten Eingabegeräten die Teilnehmer bereits Erfahrungen haben. Die Geräteklasse mit der die meisten Versuchsteilnehmer bereits Erfahrungen sammeln konnten, ist der *Bankautomat*. Bei den älteren Versuchspersonen gaben eine Teilnehmerin (71 Jahre) an, keinerlei Vorerfahrungen im Umgang mit touch-basierten Eingabegeräten zu haben. Eine weitere ältere Teilnehmerin (78 Jahre) gab an, ein touch-basiertes Mobiltelefon zu besitzen, verzichtete allerdings darauf Angaben über die Vorerfahrungen und Nutzungshäufigkeiten zu machen. Für die Berechnung der Anteile wurde diese Teilnehmerin deshalb nicht betrachtet.

5.4 AUFGABENSTELLUNGEN

Die Teilnehmer am Versuch mussten insgesamt 12 Aufgaben lösen, die sich in drei Blöcke gliedern. Der erste Block (Übungsblock) diente da-

Touch-basiertes Eingabegerät	Jüngere Probanden (n=12)	Ältere Probanden (n=11)
Mobiltelefon	100 %	36,4 %
Personal Digital Assistant (PDA)	0 %	0 %
Fahrkartenautomat	83,3 %	63,6 %
Bankautomat	100 %	81,8 %
Navigationssystem	58,3 %	18,2 %
Kopierer	83,3 %	27,3 %

Tabelle 37: Anteil der Personen mit Vorerfahrung bei der Verwendung von Geräten, die mittels Touchscreen bedient werden.

zu, dass sich die Versuchspersonen mit der Bedienung des Tablets und der Interaktion mit einem kapazitiven Touchscreens vertraut machen. Hierfür wurden zwei verschiedene Aufgabenstellungen vorgesehen.

Für die erste Aufgabe im Übungsblock musste die Versuchsperson die App *Google Maps* öffnen. Nach dem Start der App wurde der Ernst-Reuter-Platz in Berlin angezeigt. Die erste Aufgabe galt als erfüllt, wenn statt des Ernst-Reuter-Platzes der Alexanderplatz angezeigt wurde. Für die zweite Übungsaufgabe wurde die App *mMusic Mini* verwendet. Mit dieser App können Lieder zu einer Playlist hinzugefügt und abgespielt werden. Um die zweite Übungsaufgabe zu lösen musste die Versuchsperson „All you need is love“ von den Beatles erst zur Playlist hinzufügen, danach abspielen, während der Wiedergabe bis etwa zur Mitte des Liedes vorspulen und die Lautstärke der Wiedergabe verringern.

Mit den Übungsaufgaben wurden zwei Ziele verfolgt. So war ein Ziel, einen ersten Eindruck von den Vorerfahrungen der Versuchspersonen zu bekommen. Das zweite Ziel war es sicherzustellen, dass alle Versuchspersonen in der Lage sind die folgenden Testaufgaben selbstständig zu lösen. Wusste eine Versuchsperson nicht, wie eine Übungsaufgabe gelöst werden kann, so wurden der Versuchsperson verschiedene Lösungsoptionen beschrieben.

Nach Abschluss der Übungsaufgaben musste jede Versuchsperson in einem ersten Testblock insgesamt 5 verschiedene Aufgaben mit einer der in Abschnitt 5.2 beschriebenen Apps lösen. Die Entscheidung welche App als erstes verwendet wurde, ist durch das Versuchsdesign vorgegeben und wechselt zwischen zwei Versuchspersonen.

Die erste Aufgabe des Testblocks bestand darin, zwei verschiedene Bilder auf dem Fernseher anzeigen zu lassen. Für die zweite Aufgabe musste ein festgelegter Film auf dem Fernseher abgespielt werden,

zusätzlich musste in diesem Film 5 Minuten vorgespult werden. In der dritten Aufgabe ging es darum einen bestimmten TV-Kanal auf dem Fernseher abzuspielen und während der Wiedergabe die Lautstärke zu verringern. Wurde die Lautstärke verringert, musste die Versuchsperson auf einen weiteren Kanal wechseln und anschließend die Wiedergabe beenden. Die vierte Aufgabe beinhaltete den Aufruf des elektronischen Programmführers (EPG) und das Anlegen eines Timers. Ein Timer beschreibt eine geplante Aufnahme und kann direkt aus dem EPG hinzugefügt werden. Nach dem Hinzufügen des Timers musste, als fünfte Aufgabe, der Timer gelöscht werden. Für jede Aufgabe wurde der Versuchsperson vorgelesen, was zur Lösung der Aufgabe notwendig ist, allerdings wurde darauf geachtet, dass durch die Anweisungen keine Hinweise auf die Bedienelemente gegeben wurden. In der nachfolgenden Auflistung ist der Text der Aufgabe zum Anzeigen der Bilder auf dem Fernseher aufgeschrieben.

1. *Öffnen Sie die Ansicht „Bilder“.*
2. *Suchen Sie das Bild mit dem „Koala“ und lassen Sie es auf dem Fernseher anzeigen.*
3. *Wechseln Sie zum nächsten bzw. vorherigen Bild und beschreiben Sie, was Sie sehen.*
4. *Beenden Sie die Wiedergabe und schließen Sie die Ansicht auf dem Fernseher, so dass Sie zurück auf den Startbildschirm kommen.*

Der dritte Block umfasste weitere 5 Aufgabenstellungen. Die Aufgaben aus diesem Block ähneln den Aufgaben aus dem vorherigen Block, weichen jedoch in Einzelheiten ab, so wurde für die Aufgabe ein Bild auf dem Fernseher anzeigen zu lassen, in diesem Aufgabenblock, ein anderes Bild gewählt. Für die anderen Aufgabenstellungen gestalteten sich die inhaltlichen Änderungen ähnlich.

5.5 ABLAUF

Nach der Begrüßung wurden die Teilnehmer aufgefordert die Erklärung zum Testablauf und die Einverständniserklärung zur Datenverarbeitung zu lesen. Nachdem die Teilnehmer die Einverständniserklärung gelesen und unterschrieben hatten, wurden in einem ersten Fragebogen die Vorerfahrung und das Vorwissen im Umgang mit touch-basierten Eingabegeräten abgefragt. Die Vorerfahrungen wurden anhand von konkreten Beispielen (z. B. Bankautomat, Fahrkartenautomat) erhoben. Zusätzlich wurde gefragt, wie häufig die Teilnehmer mit touch-basierten Eingabegeräten interagieren. Die Spannweite der Antwortoptionen dieser Frage deckte *mehrmals täglich bis seltener als einmal pro Monat* ab.

Um das Vorwissen bezüglich touch-basierten Eingabegeräten zu ermitteln wurden die Teilnehmer zur Funktionsweise eines Touchscreens befragt. Zur Beantwortung dieser Frage standen jeder Versuchsperson drei Antwortoptionen zur Auswahl, wobei auch mehrere Optionen durch einen Teilnehmer angekreuzt werden konnten. Die erste Antwort zur Funktionsweise trifft die Aussage, dass ein Touchscreen den durch den Nutzer aufgewendeten Druck erfasst. Mit der zweiten Option wird die Aussage getroffen, dass durch den Touchscreen die Veränderung eines elektrischen Feldes gemessen wird. Für den Fall, dass die Versuchsperson die Funktionsweise eines Touchscreens nicht kennt, wurde die Antwortoption *Weiß ich nicht* aufgeführt.

Die zweite Frage zum Vorwissen, betrifft möglich Fehlerkorrekturstrategien im Umgang mit berührungsempfindlichen Touchscreens. Konkret wurde jede Versuchsperson gefragt, wie sie handeln würde, wenn die Eingabe vom Touchscreen nicht erkannt wird. Als Antwortoptionen wurden *mehr Druck aufwenden*, *Bildschirm länger berühren*, *genauer werden*, *Interaktion wiederholen* und *Weiß ich nicht* angeboten. Bei dieser Fragestellung konnten die Teilnehmer abermals mehrere Antwortoptionen ankreuzen.

Nachdem der Teilnehmer den Fragebogen zur Vorbefragung ausgefüllt hatte, wurde er oder sie mit dem verwendeten Eingabegerät, dem Nexus 7, vertraut gemacht. Allen Teilnehmern wurde erklärt wie die Eingabe über den kapazitiven Bildschirm funktioniert und welchen Zweck die Tasten auf der rechten Seite des Tablets haben. Nach der Einführung startete die erste Übungsaufgabe. Wie in Abschnitt 5.4 beschrieben mussten die Versuchspersonen zur Lösung der ersten Übungsaufgabe statt dem Ernst-Reuter-Platz den Alexanderplatz in Google Maps anzeigen lassen. Hatten die Versuchspersonen keine oder nur geringe Vorkenntnisse, so wurden sie erst aufgefordert etwas zu versuchen, bevor ihnen erläutert wurde, wie sie interagieren können, um die Aufgabe zu lösen. Zeigte sich während der Übungsaufgabe ein Fehler, so wurde der Versuchsperson erklärt welcher Fehler aufgetreten war und wie er vermieden werden kann. Nachdem die Versuchsperson die Übungsaufgaben gelöst hatte, wurden ihr zur Bewertung zwei Fragebögen vorgelegt. Der erste Fragebogen erfasste, wie anstrengend die gestellte Aufgabe war. Zur Beantwortung dieser Frage, musste die Versuchsperson eine Markierung auf der SEA-Skala (Skala zur Erfassung subjektiv erlebter Anstrengung [58]) machen. Zur Bewertung der Attraktivität und pragmatischen Qualität der App, wurde der AttrakDiff Mini von Hassenzahl [88] verwendet.

Nachdem die Übungsaufgaben von der Versuchsperson gelöst wurden, wurde die erste App zur Steuerung von XBMC durch die Versuchsperson gestartet. Bevor die entsprechenden Aufgaben vom Teilnehmer mit der App bearbeitet wurden, wurde mit der Versuchsperson die Oberflächen der App und des Mediacenters besprochen. Da die originale XBMC-Remote nur in einer englischen Version ver-

füßbar ist (vgl. Abschnitt 5.2) wurde jeder Menüpunkt angesprochen und das Label des Buttons übersetzt. Für die Yatse-App wurden alle Piktogramme erklärt. Wurden vom Teilnehmer keine weiteren Fragen zur App oder zur Bedienung der Oberfläche des Mediacenters gestellt, so wurden die in Abschnitt 5.4 beschriebenen Aufgaben der Versuchsperson vorgelesen. Nach jeder Aufgabe wurde dem Versuchsteilnehmer die SEA-Skala vorgelegt, um die gerade abgeschlossene Aufgabe hinsichtlich der erlebten Anstrengung zu bewerten. Am Ende der 5 Testaufgaben wurde der Versuchsperson zusätzlich der AttrakDiff Mini vorgelegt.

Bevor die zweite App zur Anwendung kam, wurden dem Teilnehmer zwei Fragebögen zur Erfassung der Technikaffinität und zur Technikeinstellung vorgelegt. Zur Erfassung der Technikaffinität wurde wiederum der TA-EG Fragebogen von Karrer et al. [127] verwendet. Die Technikeinstellung wurde mit dem ICT Fragebogen von Weiss et al. [247] erfasst. Nach dem Ausfüllen der Fragebögen begann der Test der zweiten App. Erneut wurde vor der Versuchsdurchführung die Oberfläche der verwendeten App mit dem Teilnehmer besprochen, um Probleme aufgrund von fehlenden Sprachkenntnissen oder unklaren Piktogrammen zu vermeiden. Wie schon im ersten Testblock, wurde auch in diesem Testblock dem Teilnehmer nach jeder Aufgabe eine SEA-Skala zur Bewertung der subjektiven Anstrengung und am Ende der Durchführung der AttrikDiff Mini zur Bewertung der Attraktivität und pragmatischen Qualität vorgelegt.

Den Abschluss eines jeden Versuches bildete der demografische Fragebogen. Mit dem demografischen Fragebogen, wurden das Geschlecht, das Alter, der höchste Schulabschluss, die beruflichen Abschlüsse und der aktuelle Beschäftigungsstand abgefragt. Zusätzlich wurde durch den demografischen Fragebogen erhoben, wie häufig die Versuchsperson zu Hause für private Zwecke das Internet nutzt und ob die Personen einen Fernseher mit Internetfunktion, ein sogenanntes Smart-TV, besitzt.

5.6 ERGEBNISSE

Nachfolgend werden die Ergebnisse aus dem empirischen Versuch vorgestellt. Neben den reinen Interaktionsparametern werden auch die Nutzerattribute der Versuchspersonen betrachtet und mit den Daten aus Kapitel 4 verglichen. Nach den empirischen Ergebnissen werden die Ergebnisse der Simulation und die vorhergesagten Nutzermodelle dargestellt und erläutert. Weiterhin werden die erzielten Ergebnisse mit den Ergebnissen des empirischen Versuchs verglichen und deren Vorhersagegüte beurteilt.

Auslöser	Jüngere Probanden (n=12)	Ältere Probanden (n=12)
Überschreitung des Mindestdrucks	16,7 %	0 %
Änderung im elektri- schen Feld	25 %	25 %
Beides	8,3 %	8,3 %
Unbekannt	50 %	66,7 %

Tabelle 38: Auslöser für die Eingabeerkennung eines kapazitiven Touchscreens

5.6.1 Empirischer Versuch

Bereits in den Übungsaufgaben zeigte sich, dass alle Versuchsteilnehmer eine Vorstellung davon hatten, wie das Tablet bedient werden muss. Alle Teilnehmer wussten bzw. vermuteten, dass das Tablet durch die Berührung des Bildschirms mit dem Finger bedient werden kann. Es zeigten sich allerdings schon erste Unterschiede zwischen den Versuchspersonengruppen, was die Ausführung der Interaktion (mit oder ohne Druck; längeres oder kürzeres Berühren) betraf.

In Tabelle 38 ist dargestellt, wie die Versuchsteilnehmer sich vorstellen, wie eine Eingabe auf einem modernen Touchscreen ausgelöst wird. Interessant ist, dass in beiden Gruppen die Mehrheit nicht weiß wie eine Aktion ausgelöst werden kann bzw. dass die Probanden sich nicht festlegen wollten. Ein Viertel der jüngeren und älteren Versuchspersonen wusste bzw. vermutete, dass moderne berührungsempfindliche Bildschirme, wie sie in Smartphones eingebaut werden, die Veränderung des elektrischen Feldes messen. Je eine ältere und eine jüngere Versuchsperson gab an, dass beide Arten in Mobiltelefonen verbaut werden. Zwei jüngere Versuchspersonen gaben an, dass die Eingabe auf dem Bildschirm des Smartphones über den aufgebrachten Druck auf den Bildschirm ausgelöst wird.

In Tabelle 39 sind die Korrekturstrategien nach einem Erkennungsfehler dargestellt, wie sie von den Versuchspersonen angewendet werden würden. Wie der Tabelle zu entnehmen ist, würde die Hälfte der jüngeren Versuchspersonen als Fehlerkorrekturstrategie die Eingabe entweder *einfach wiederholen* oder *bei der Ausführung genauer werden*. Bei den älteren Versuchspersonen ist das Bild nicht so eindeutig. Ein Drittel der älteren Versuchspersonen würde als Korrekturstrategie die *Eingabe einfach wiederholen*. Ebenso viele Probanden wissen nicht, was sie tun können, wenn der Touchscreen die Eingabe nicht erkannt hat.

Strategie(n) zur Fehlerkorrektur	Jüngere Probanden (n=12)	Ältere Probanden (n=12)
Nur mehr Druck aufwenden	1	0
Mehr Druck aufwenden bzw. Bildschirm länger berühren	1	1
Bildschirm länger berühren bzw. Eingabe wiederholen	0	2
Bildschirm länger berühren, Eingabe wiederholen bzw. Genauer werden	1	0
Nur genauer werden	0	1
Genauer werden bzw. Eingabe wiederholen	6	0
Nur die Eingabe wiederholen	1	4
Alle Strategien sind möglich	1	0
Weiß nicht	1	4

Tabelle 39: Korrekturstrategie für den Fall, dass der Touchscreen die Eingabe nicht erkannt hat

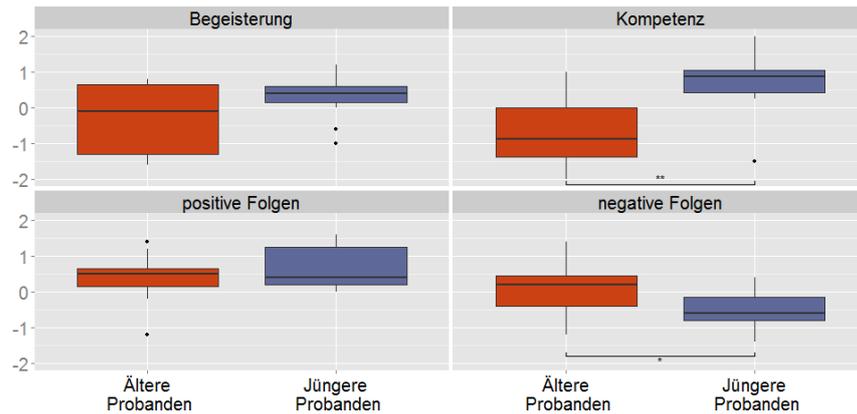


Abbildung 47: Verteilung in den Dimensionen der Technikaffinität für die 2 Versuchspersonengruppen

Bei den jüngeren Probanden wusste ein Mann nicht, wie er in diesem Fall reagieren würde.

5.6.1.1 Technikaffinität und Technikeinstellung der Versuchspersonen

In diesem Unterabschnitt werden die Ergebnisse zur Technikaffinität und Technikeinstellung in den verschiedenen Altersgruppen präsentiert. Da in der Gruppe der älteren Versuchspersonen das Verhältnis zwischen den Geschlechtern unausgewogen ist (11 Frauen : 1 Mann), werden die Nutzerattribute nur anhand der Altersgruppen verglichen.

In der Abbildung 47 sind die Verteilungen für die Dimensionen der Technikaffinität, in den beiden Altersgruppen, dargestellt. Wie in der Abbildung kenntlich gemacht, existieren signifikante Unterschiede zwischen jüngeren und älteren Versuchspersonen in den Dimensionen *Kompetenz* und *negative Einschätzung von Technikfolgen*. Die älteren Versuchspersonen schätzen ihre Kompetenz im Umgang mit Technik signifikant niedriger ($t_{(22)} = -3,74; p = 0,001$) ein, als dies die jüngeren Teilnehmer tun. Hinsichtlich der negativen Technikfolgenabschätzung zeigen die Senioren signifikant höhere Werte ($t_{(22)} = 2,41; p = 0,026$), als die jüngeren Versuchspersonen. Für die Dimensionen *positive Einschätzung der Technikfolgen* und *Begeisterung für Technik* existieren keine signifikanten Unterschiede zwischen den Gruppen.

Die Abbildung 48 zeigt die Verteilungen für die Dimensionen der Technikeinstellung, für die verschiedenen Altersgruppen. Bis auf die Design-Dimension zeigen alle Dimensionen signifikante Unterschiede zwischen älteren und jüngeren Versuchspersonen. Die Ergebnisse des Vergleichs sind für die zweiseitigen T-Tests in Tabelle 40 aufgeführt. Da in der Dimension *Interesse an Technik* die Verteilung der Varianzen (Homoskedastizität) zwischen den Gruppen signifikant unterschiedlich ist, wurde der zweiseitige Welch-Test als Alternative zum T-Test verwendet.

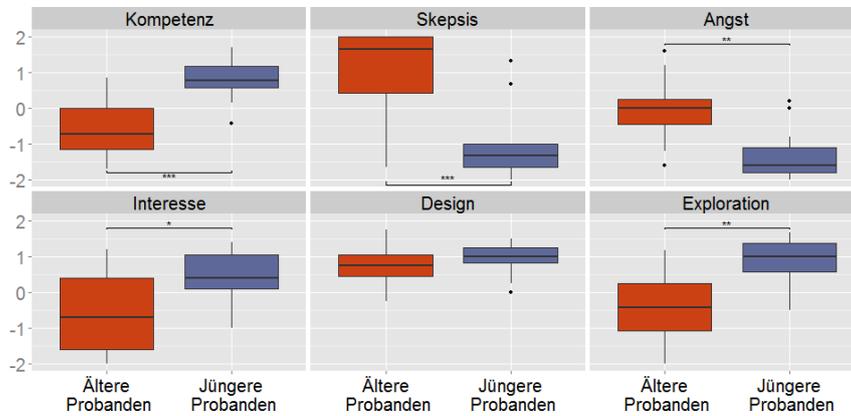


Abbildung 48: Verteilung in den Dimensionen der Technikeinstellung (ICT) für die 2 Versuchspersonengruppen

Dimension	$t_{(22)}$	p
Kompetenz	-4,558	< 0,001
Skepsis	4.212	< 0,001
Angst	3.58	0.002
Interesse	-2.433	0.025
Exploration	-3.877	0,001

Tabelle 40: T-Werte und Signifikanzniveaus für die Dimensionen *Kompetenz*, *Skepsis*, *Angst*, *Interesse* und *Exploration*

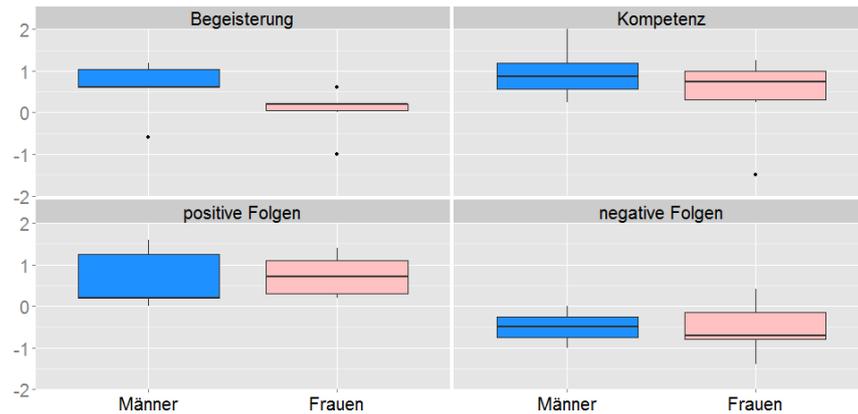


Abbildung 49: Verteilung in den Dimensionen der Technikaffinität für die jüngeren Frauen und Männer

Die in Abbildung 48 dargestellten Verteilungen zeigen stellenweise deutliche Unterschiede zwischen den beiden Altersgruppen. Besonders groß ist der Unterschied in der Dimension *Skepsis*, die das Misstrauen einer Person hinsichtlich finanzieller Aktivitäten unter Verwendung von Informations- und Kommunikationstechnik (IKT) ausdrückt. In dieser Dimension haben die älteren Versuchspersonen einen Mittelwert von 1,03 (SD = 1,37) und die jüngeren Teilnehmer einen Mittelwert von -1,056 (SD = 1,023). Bei einer Skala von -2 bis +2, drücken Werte über Null aus, dass die IKT eher nicht geeignet ist, um finanzielle Aktivitäten darüber abzuwickeln. Wie an den Werten abzulesen ist, sind die Einschätzungen der Altersgruppen, in dieser Dimension, konträr zueinander.

Um zu überprüfen, ob die gefundenen Unterschiede zwischen den Altersgruppen oder den Geschlechtern existieren, werden nachfolgend die Verteilungen für die Technikaffinität und Technikeinstellung in der Gruppe der jüngeren Versuchsteilnehmer untersucht. Die Abbildung 49 sind die Verteilungen für die Dimensionen der Technikaffinität, in der Gruppe der jüngeren Versuchsteilnehmer, dargestellt. Die Überprüfung der Unterschiede zwischen den Geschlechtern, für die einzelnen Dimensionen, ergibt, was sich bereits in der Abbildung andeutet, dass keine signifikanten Unterschiede zwischen den Gruppen bestehen. So sind die p -Werte für alle Signifikanztests größer 0,05.

In Abbildung 50 sind die Verteilungen für die Dimensionen der Technikeinstellung bei den jüngeren Männern und Frauen zu sehen. Wie bereits bei der Technikaffinität existieren auch in den Dimensionen der Technikeinstellung keine signifikanten Unterschiede zwischen den Geschlechtern. Die in diesem Abschnitt gefundenen Unterschiede hinsichtlich der Technikaffinität und Technikeinstellung sind somit, mit großer Wahrscheinlichkeit, auf die Altersgruppen zurückzuführen.

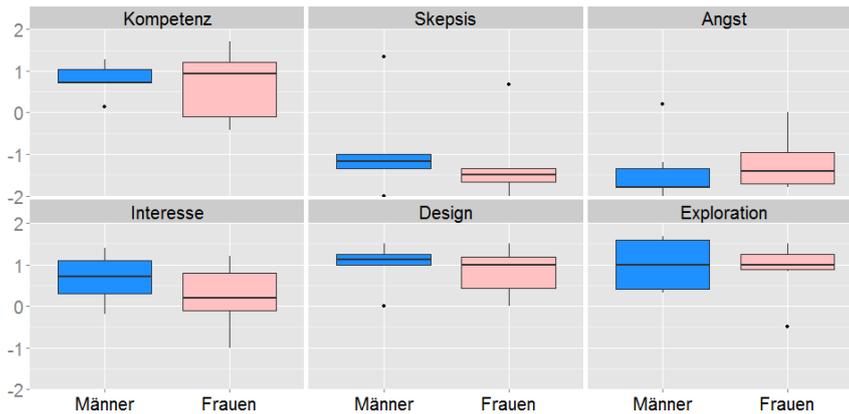


Abbildung 50: Verteilung in den Dimensionen der Technikeinstellung (ICT) für die jüngeren Frauen und Männer

5.6.1.2 Interaktionsparameter

In diesem Unterabschnitt werden die Interaktionsparameter aus dem Versuch dargestellt. Der Abschnitt beginnt mit der allgemeinen Darstellung der benötigten Interaktionsschritte. Im weiteren Verlauf werden die Unterschiede im Interaktionsverhalten in den verschiedenen Altersgruppen detailliert erläutert, um einen Einblick in das Interaktionsverhalten und die daraus resultierenden Probleme zu geben.

In Abbildung 51 sind die, für die Lösung aller Aufgaben, benötigten Schritte dargestellt. Wie in der Darstellung angezeigt, existieren zwischen den jüngeren und älteren Personen signifikante Unterschiede ($U = 124$; $p = 0,002$) hinsichtlich der Anzahl der angewendeten Interaktionen. Die älteren Versuchspersonen brauchten im Mittel 349 Interaktionen ($M = 349,417$; $SD = 61,317$), um alle Aufgaben zu lösen, wohingegen die jüngeren Versuchsteilnehmer nur 275 Interaktionen ($M = 275,167$; $SD = 90,101$) benötigten.

Die Abbildung 52 stellt die Verteilung der Interaktionstypen dar, wie sie vom System in den verschiedenen Altersgruppen erkannt wurden. Wie aus der Grafik ersichtlich wird, ist der *Short tap* die Interaktionsform, die am häufigsten vom System erkannt wurde. Ein *Short tap* ist, in diesem Versuch, wie folgt definiert: Die Berührung des Touchscreen dauert nicht länger als 250 ms und die Distanz zwischen Aufsetzpunkt und Abhebepunkt beträgt nicht mehr als 50 Pixel. Die Festlegung der Zeit erfolgte anhand der Logfiles des verwendeten Tablets. Die maximale Distanz zwischen Aufsetzpunkt und Abhebepunkt resultiert aus der im Android Quellcode beschriebenen Toleranz bezüglich der maximalen Abweichung [6]. Bei den jüngeren Teilnehmern wurde eine Interaktion in rund 81 % der Fälle ($M = 80,574$; $SD = 4,12$) vom System als *Short tap* interpretiert, während die älteren Teilnehmer in rund 71 % der Fälle ($M = 71,488$; $SD = 11,583$) einen *Short tap* verwendeten. Der T-Test zeigt, dass die jüngeren Versuchsteilnehmer signifikant

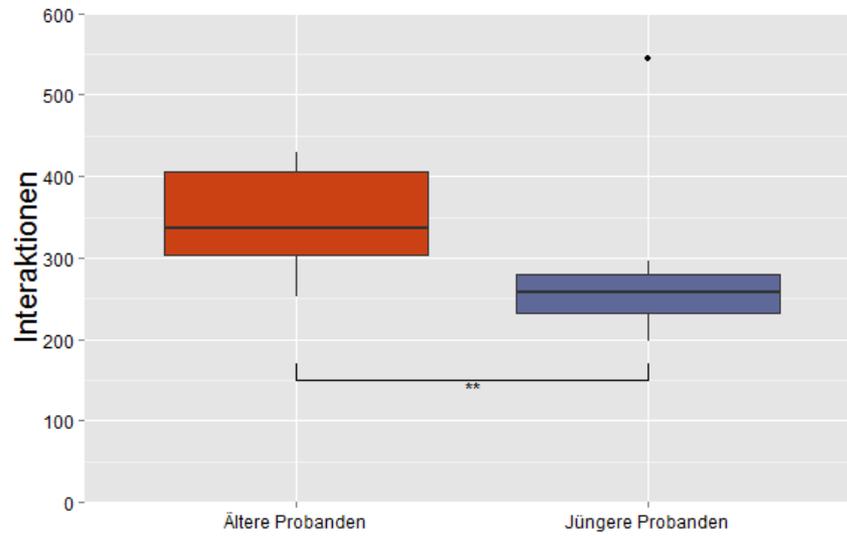


Abbildung 51: Verteilung für die benötigten Interaktionen in den zwei Altersgruppen

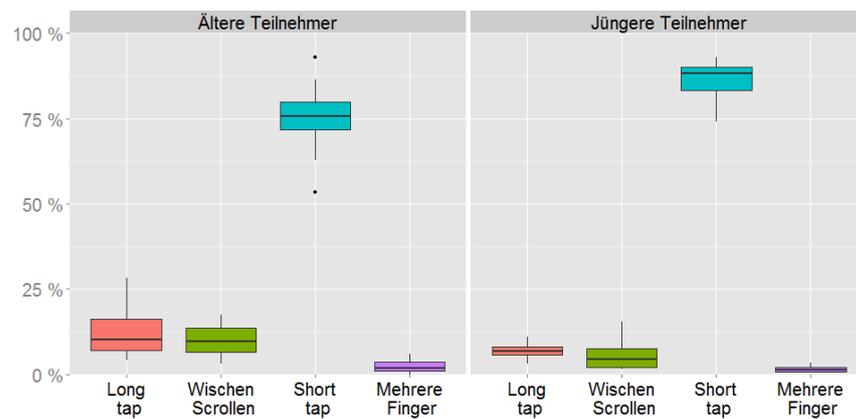


Abbildung 52: Verteilung für die vom System erkannten Interaktionstypen in den zwei Altersgruppen

($t_{(22)} = 2,56; p = 0,023$) häufiger Interaktionen verwenden, die als *Short tap* erkannt werden.

Für die Erkennung von *Long tap*, *Swipe* (Scrollen oder Wischen) und Interaktionen mit *mehreren Fingern* auf dem Bildschirm ändert sich das Bild. Beim *Long tap* zeigt der zweiseitige Welch-Test keine signifikanten Unterschiede ($t_{(22)} = 2,149; p > 0,05$) zwischen den älteren Versuchsteilnehmer ($M = 15,138; SD = 7,359$) und jüngeren Teilnehmer ($M = 10,314; SD = 2,518$) an. Ähnliches findet sich beim Scrollen und Wischen über den Touchscreen. Auch hier existieren keine signifikanten Unterschiede ($t_{(22)} = 1,903; p > 0,05$) zwischen den Altersgruppen. Die älteren Teilnehmer ($M = 8,961; SD = 4,806$) führen ebenso viele Eingaben aus, die vom System als Scrollen/Wischen interpretiert werden, wie dies die jüngeren Teilnehmer ($M = 5,997; SD = 2,455$) tun. Auch für die Verwendung von mehreren Fingern (Kontaktpunkten) auf dem Touchscreen existieren keine signifikanten Unterschiede zwischen den Altersgruppen ($t_{(22)} = 1,282; p > 0,05$). Bei den älteren Versuchspersonen sind etwa 4 % ($M = 4,413; SD = 3,169$) und bei den jüngeren Versuchspersonen in etwa 3 % ($M = 3,115; SD = 1,504$) der Interaktionen mit mehr als einem Kontaktpunkt auf dem Touchscreen ausgeführt worden.

In den Abbildungen 53 und 54 sind die im empirischen Versuch beobachteten Interaktionen, nach Zeit und Distanz zwischen Aufsetzpunkt und Abhebepunkt getrennt, für die älteren und jüngeren Teilnehmer dargestellt. Die Dauer repräsentiert die Zeit in Millisekunden, die (mindestens) ein Finger den Touchscreen berührte. Die Distanz stellt für die Interaktionsformen *Short tap*, *Long tap* und *Scrollen/Wischen* den maximalen Abstand zwischen Aufsetzpunkt und Abhebepunkt dar. Für die Interaktion mit mehreren Aufsetzpunkten (*Mehrere Finger*), stellt die Distanz den maximalen Abstand zwischen den verschiedenen Berührungspunkten dar. Werden die Abbildungen verglichen, so fällt auf, dass die Streuung bei den älteren Teilnehmern (Abbildung 53) deutlich größer ist, als bei den jüngeren Teilnehmern (Abbildung 54). Eine höhere Streuung bedeutet, dass die Interaktionen der älteren Versuchsteilnehmer heterogener als die der jüngeren Teilnehmer sind.

Ein ähnliches Bild zeigt sich, wenn die vergrößerten Ausschnitte in der Abbildungen 55 und 56 betrachtet werden. Wie in der Abbildung 55 zu sehen ist, streut die Distanz zwischen Aufsetzpunkt und Abhebepunkt stark in der Gruppe der älteren Teilnehmer. Ist die Distanz zwischen Aufsetzpunkt und Abhebepunkt größer als 50 Pixel, so wird die Interaktion vom Tablet als Scroll- bzw. Wischgeste interpretiert. Bei einer Auflösung von 1280 x 800 Pixel und einer Displaygröße von 15,1 x 9,45 cm ergibt sich somit eine Distanz von rund 0,6 cm, ab der das System die Eingabe als Scroll- bzw. Wischgeste verarbeiten wird. In den Abbildungen 55 und 56 ist der Bereich zwischen 50 und 100 Pixel mit Orange hervorgehoben. Wie in Abbildung 55 zu sehen ist,

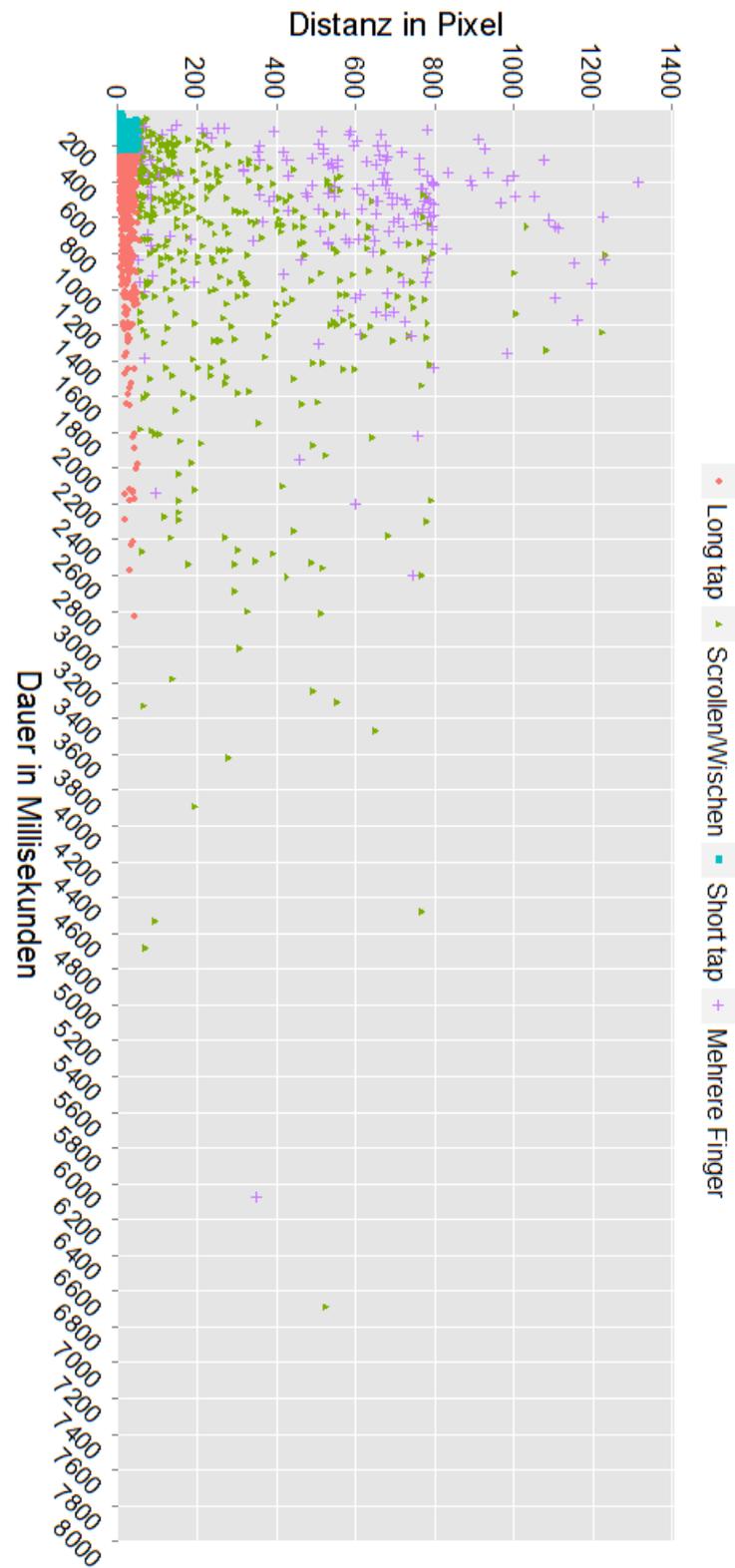


Abbildung 53: Interaktionen im empirischen Test bei den älteren Versuchspersonen

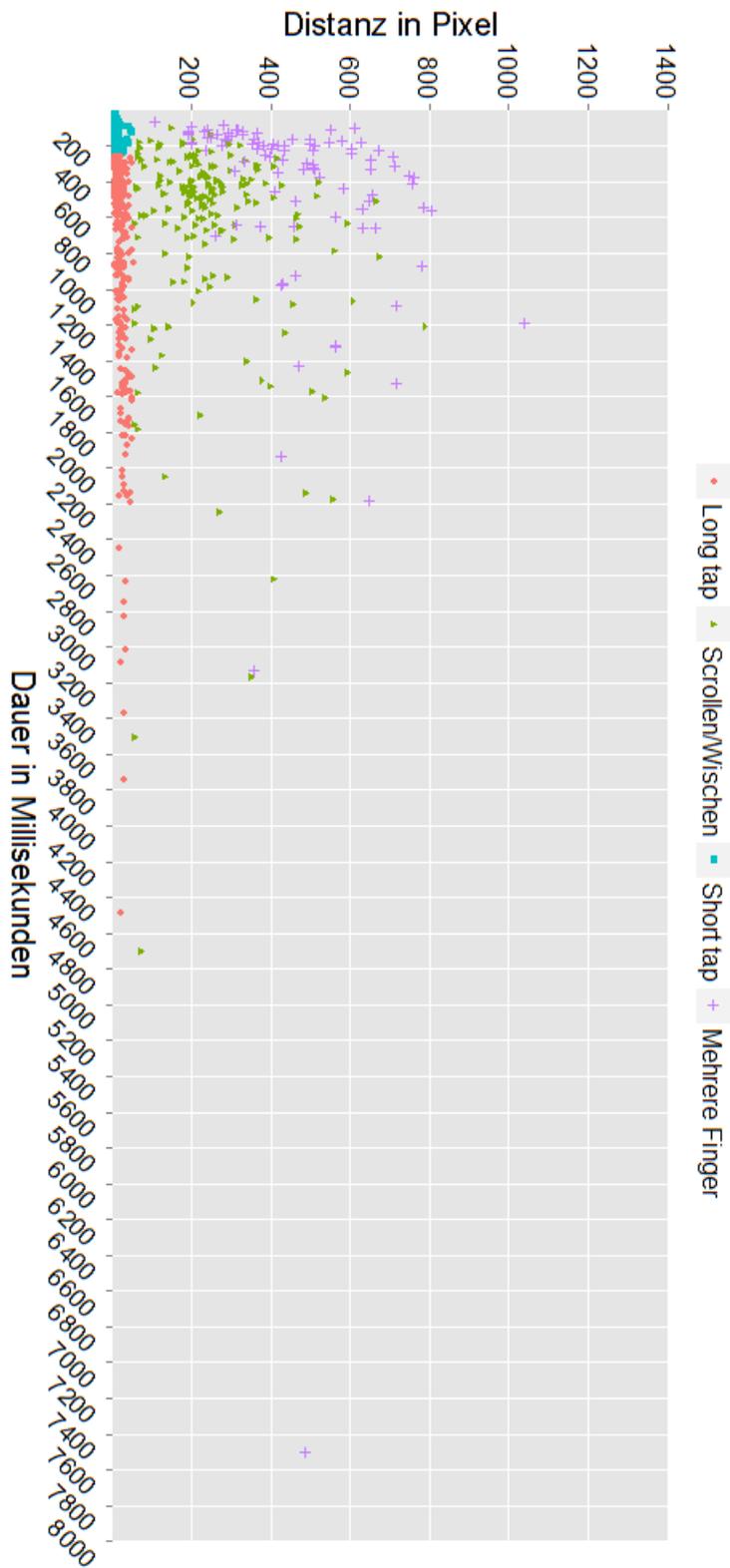


Abbildung 54: Interaktionen im empirischen Test bei den jüngeren Versuchspersonen

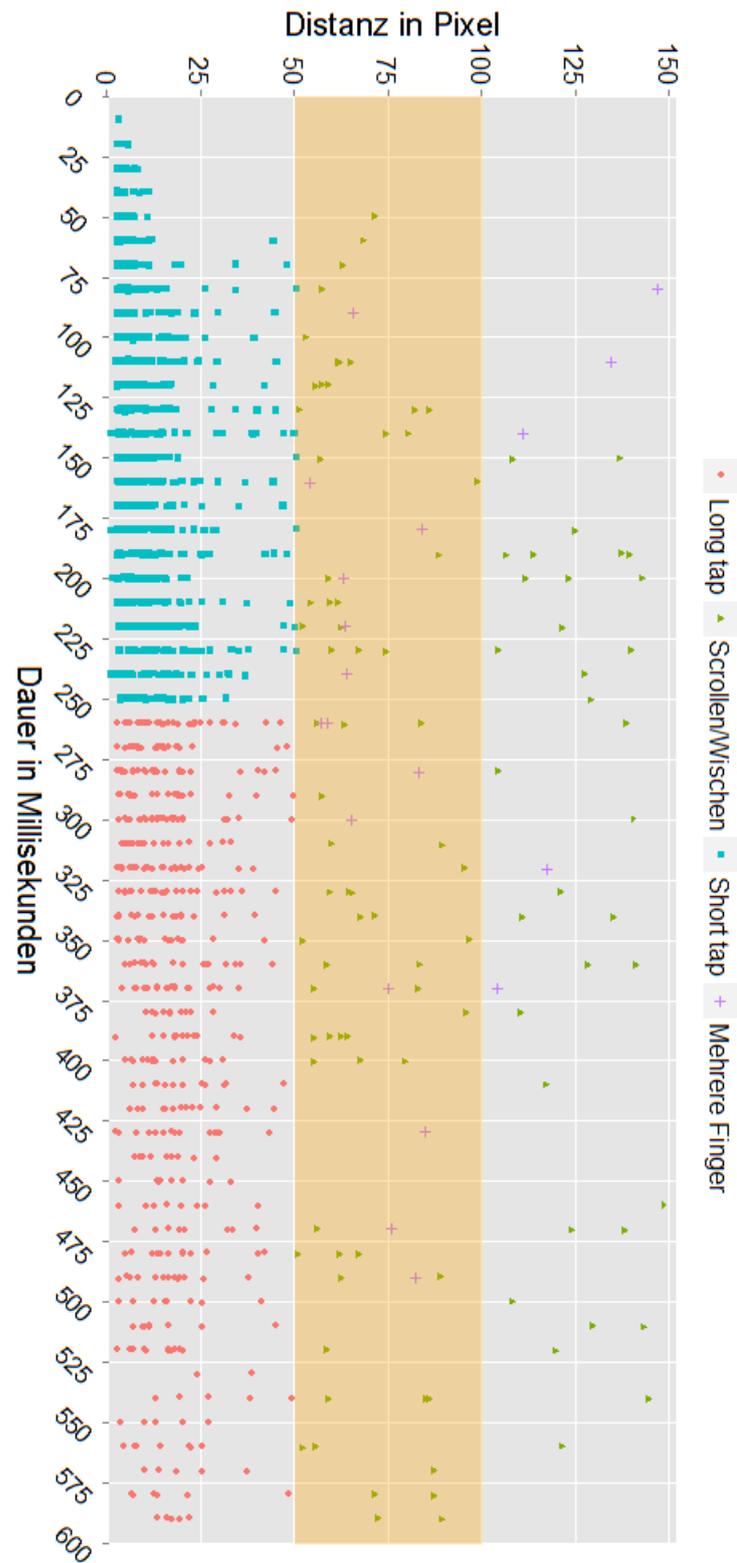


Abbildung 55: Interaktionen im empirischen Test bei den älteren Versuchspersonen (Distanz < 150 px & Dauer < 600 ms)

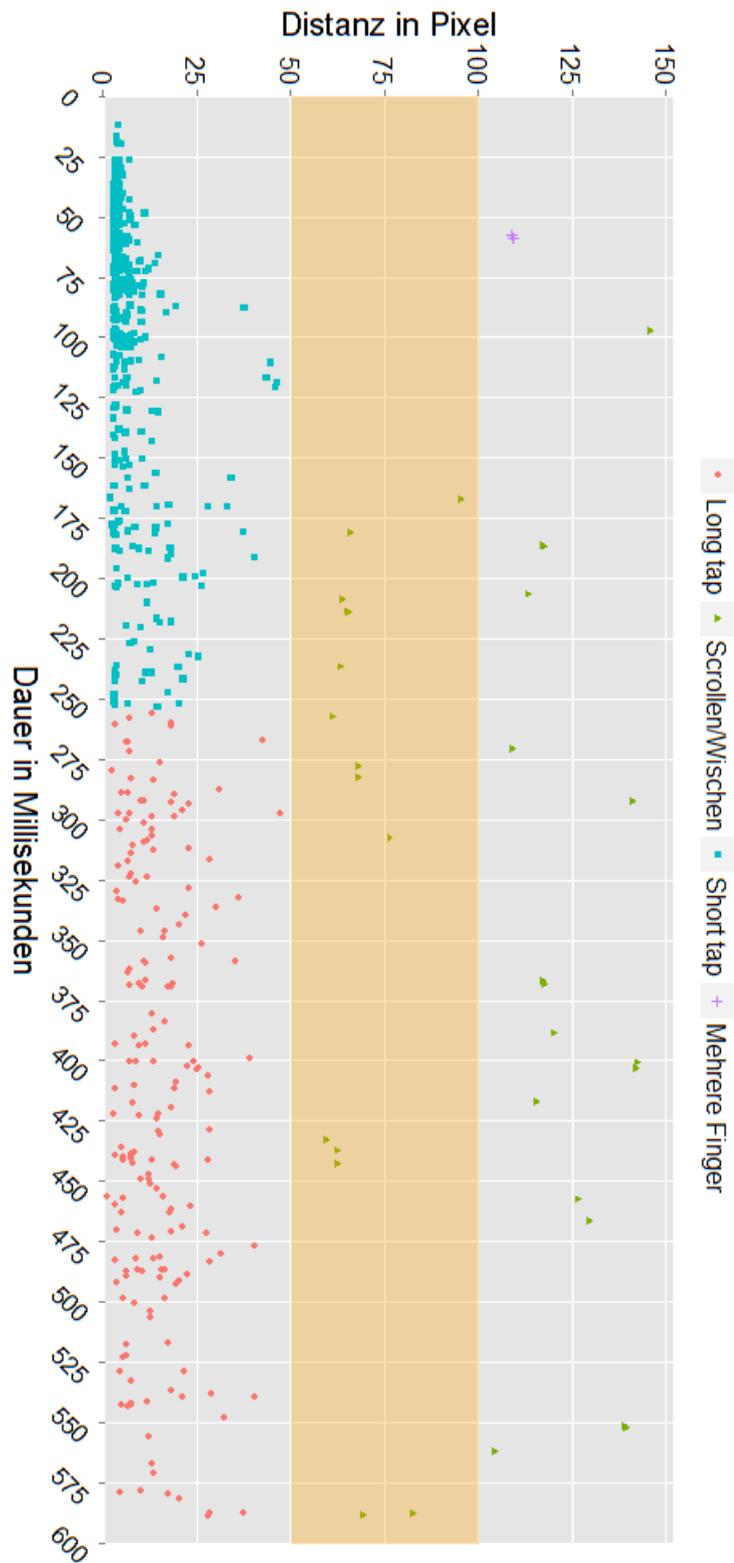


Abbildung 56: Interaktionen im empirischen Test bei den jüngeren Versuchspersonen (Distanz < 150 px & Dauer < 600 ms)

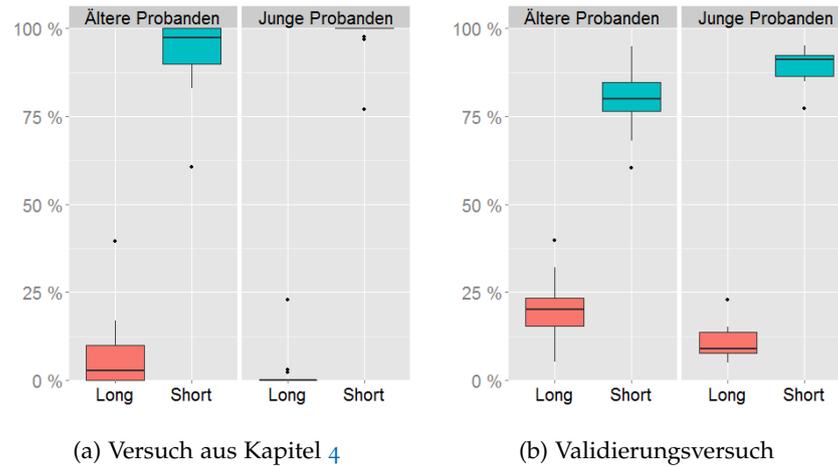


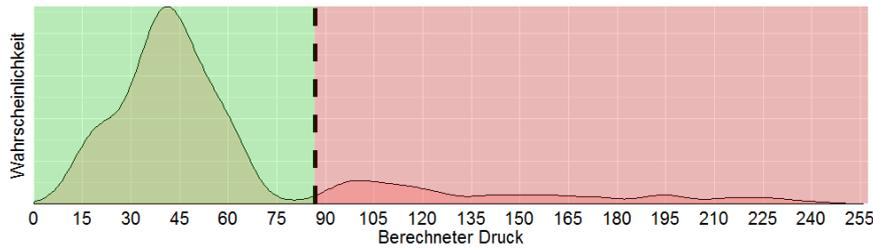
Abbildung 57: Vergleich der Interaktionsdauer in den Altersgruppen für den Validierungsversuch und den Versuch aus Kapitel 4

wurden, für die Gruppe der älteren Versuchsteilnehmer, vom Tablet viele Interaktionen ($n = 123$; 2,93 %) erfasst, die eine maximale Distanz zwischen 50 und 100 Pixel hatten. Für die Gruppe der jüngeren Teilnehmer (Abbildungen 56) ist der Wert geringer ($n = 26$; 0,79 %). Wird angenommen, dass mit den Interaktionen nicht gescrollt/gewischt werden sollte, so kann hierdurch bereits ein Teil der zusätzlich nötigen Interaktionen, in der Gruppe der älteren Versuchsteilnehmer, erklärt werden.

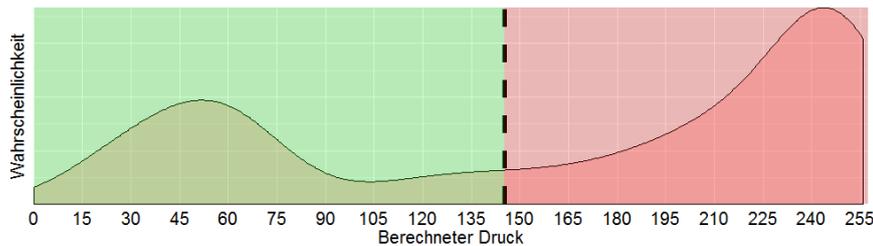
Um die Ergebnisse des aktuellen Versuchs mit den Ergebnissen des vorherigen Versuchs aus dem Kapitel 4 vergleichen zu können, werden die Interaktionen der Versuchsteilnehmer hinsichtlich der Interaktionsdauer und des aufgewendeten Drucks betrachtet.

In Abbildung 57 sind die Anteile der langen und kurzen Interaktionen in den Altersgruppen für die beiden Versuche dargestellt. Wird die Abbildung für den Versuch aus Kapitel 4 (Abbildung 57a) mit der Abbildung 57b (Validierungsversuch) verglichen, so fällt auf, dass sowohl die jüngeren als auch die älteren Probanden im Validierungsversuch häufiger längere Interaktionen verwendet haben. Hieraus kann der Schluss gezogen werden, dass die Aufgabenstellung bzw. die App im Validierungsversuch ein anderes Interaktionsverhalten verlangt. Ebenfalls auffällig ist, dass in beiden Versuchen der Anteil der langen Interaktionen in der Gruppe der älteren Versuchspersonen höher ist. Werden die Verteilungen in den Abbildungen 55 und 56 betrachtet, so kann festgehalten werden, dass ältere Menschen in ihren Interaktionen mehr Zeit brauchen und nicht immer eine lange Interaktion beabsichtigen.

Für die Unterscheidung, ob ein Nutzer für eine Interaktion *Druck* oder *keinen Druck* verwendet, wurde mittels der *density*-Funktion in R eine Kerndichteschätzung für die vom System erfassten Werte eines



(a) Jüngere Versuchsperson



(b) Ältere Versuchsperson

Abbildung 58: Kerndichteschätzung für die vom System berechneten Werte für den aufgetragenen Druck, inkl. Unterscheidung zwischen Touch (grün) und Press (rot)

Nutzers berechnet. Eine Kerndichteschätzung dient nach Silverman [216] zur Schätzung der wahrscheinlichen Verteilung einer Zufallsvariable.

Die Abbildung 58 sind als Beispiel je eine Kerndichteschätzung für eine jüngere Versuchsperson (Abbildung 58a) und einen älteren Versuchsteilnehmer (Abbildung 58b) dargestellt. Beide Abbildungen weichen von einer Normalverteilung ab, was auf unterschiedliche Interaktionen bzgl. des aufgetragenen Drucks schließen lässt. Die in den Abbildungen mit grün markierten Bereiche enthalten die Werte, die als *Touch* gewertet werden sollen. Im roten Bereich sind die Werte enthalten, die als *Press* klassifiziert werden sollen. Die Berechnung der Bereiche erfolgt wie folgt.

In einem ersten Schritt werden der Mittelwert (M_{Druck}) und die Standardabweichung (SD_{Druck}) der vom System erfassten Druckwerte pro Versuchsperson berechnet. Die Berechnung der beiden Werte erfolgt, um den Einfluss der Fingergröße des Probanden und unwillkürlichen Schwankungen im aufgetragenen Druck zu minimieren. Anschließend wird mittels der *density*-Funktion die Kerndichteschätzung durchgeführt, die für jeden x-Wert (Druck) einen y-Wert (Häufigkeit) schätzt. Der x-Wert mit der größten Wahrscheinlichkeit (dem höchsten y-Wert), der kleiner oder gleich dem Mittelwert (M_{Druck}) ist, soll dem Druck-Normalwert (p_{Normal}) für die Versuchsperson entsprechen.

Zur Berechnung der Barriere, ab der eine Interaktion als *Press* und nicht mehr als *Touch* gewertet wird, werden der Druck-Normalwert (p_{Normal}) der Versuchsperson mit der Standardabweichung (SD_{Druck})

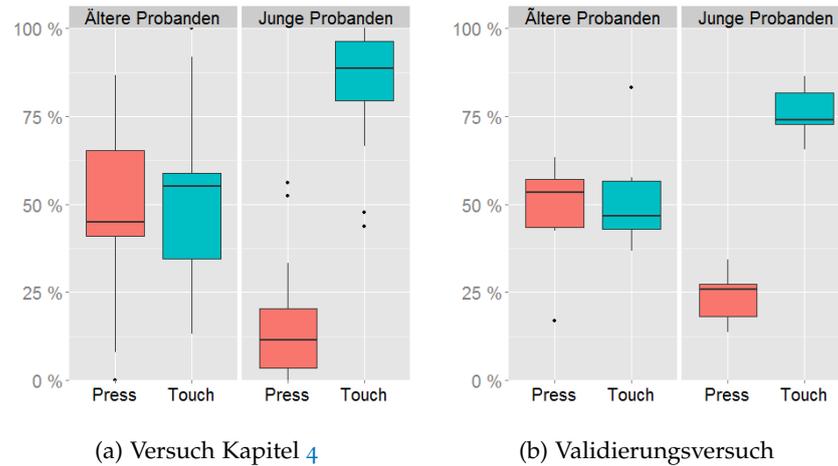


Abbildung 59: Anteil von *Press*- und *Touch*-Interaktionen für ältere und jüngere Versuchspersonen in zwei Experimenten

der Versuchsperson addiert. Durch die Addition der Standardabweichung wird sichergestellt, dass unwillkürliche Schwankungen im durch die Versuchsperson aufgebrauchten Druck ausgeglichen werden und dass die Klassifikation des Drucks, in *Touch* oder *Press*, nicht durch kleine Abweichungen beeinflusst wird.

In Tabelle 41 sind die Barrieren und Anteile für *Touch*-Interaktionen für die einzelnen Versuchspersonen aufgeführt. Für die älteren und jüngeren Teilnehmer sind die Versuchspersonen hervorgehoben, deren *Touch*-Anteil am niedrigsten (rot) bzw. am höchsten (grün) ist. Werden die Werte der beiden Gruppen verglichen, so fällt auf, dass der *Touch*-Anteil für die ältere Versuchspersonengruppe mit einem durchschnittlichen Wert von 49,92 % (SD = 13,542) signifikant geringer ($U = 11$; $p < 0,001$) ist, als der *Touch*-Anteil bei den jüngeren Versuchspersonen ($M = 76,38$ %; $SD = 5,95$). Bei den *Touch/Press*-Barrieren sind die durchschnittlichen Werte bei den älteren Versuchspersonen $M = 77,667$; $SD = 35,525$ und bei den jüngeren Teilnehmern $M = 53,25$; $SD = 20,574$. Der gefundene Unterschied ist nicht signifikant ($U = 101$; $p = 0,1$).

In Abbildung 59 sind die Verteilungen für *Touch* und *Press*, in den beiden Versuchspersonengruppen, für den Validierungsversuch (Abbildung 59b) und den Versuch aus dem Kapitel 4 (Abbildung 59a), dargestellt. Wie schon in dem im Kapitel 4 beschriebenen Versuch, wird auch im Validierungsversuch deutlich, dass sich das Interaktionsverhalten von jüngeren und älteren Personen hinsichtlich des aufgewendeten Drucks unterscheidet. In beiden Versuchen wird bei den älteren Teilnehmern rund die Hälfte der Interaktionen mit Druck ausgeführt. Wird der *Touch*-Anteil der älteren Teilnehmer am Versuch aus Kapitel 4 ($M = 52,3$ %; $SD = 25,55$) mit dem Anteil der *Touch*-Interaktionen der ältere Teilnehmer am Validierungsversuch ($M =$

Versuchsperson	Alter	Geschlecht	Barriere	Anteil <i>Touch</i>
VP 01	73	weiblich	141	36,48 %
VP 02	78	weiblich	109	46,76 %
VP 03	70	weiblich	49	59,24 %
VP 04	80	weiblich	71	57,52 %
VP 05	68	weiblich	146	37,05 %
VP 06	71	männlich	65	38,33 %
VP 07	70	weiblich	64	43,53 %
VP 08	70	weiblich	37	84,96 %
VP 09	72	weiblich	75	45,94 %
VP 10	72	weiblich	68	43,53 %
VP 11	72	weiblich	47	57,3 %
VP 12	75	weiblich	60	48,4 %
VP 13	21	männlich	32	81,79 %
VP 14	20	männlich	56	74,71 %
VP 15	30	männlich	88	81,4 %
VP 16	18	männlich	43	71,1 %
VP 17	24	männlich	81	73,85 %
VP 18	28	männlich	72	71,86 %
VP 19	26	weiblich	74	77,44 %
VP 20	29	weiblich	38	73,28 %
VP 21	27	weiblich	38	75,87 %
VP 22	26	weiblich	34	86,03 %
VP 23	28	weiblich	52	65,49 %
VP 24	28	weiblich	31	83,78 %

Tabelle 41: Barrieren und Anteile der *Press-* bzw. *Touch*-Interaktionen pro Versuchsperson

49,92 %; SD = 13,542) verglichen, so zeigt sich kein signifikanter Unterschied ($U = 74$; $p > 0,05$) zwischen den Versuchen. Für die jüngeren Versuchsteilnehmer ist der Unterschied zwischen den Ergebnissen aus dem Kapitel 4 ($M = 82,97$ %; $SD = 18,37$) und den Ergebnissen aus dem Validierungsversuch ($M = 76,38$ %; $SD = 5,95$) ebenfalls nicht signifikant ($t_{(24)} = 1,266$; $p > 0,05$).

Für den Abschnitt über die Interaktionsparameter kann zusammenfassend festgehalten werden: Die Interaktionen von älteren Versuchspersonen dauern meist länger und bei der Interaktion wird oft mehr Druck ausgeübt. Aus beiden Abweichungen können Fehler entstehen, die die Interaktion mit dem System beeinflussen.

5.6.1.3 Zwischenfazit

Bezüglich der in diesem Versuch ermittelten Nutzerattribute ist anzumerken, dass die Unausgewogenheit der Geschlechterverteilung für die Gruppe der älteren Versuchsteilnehmer als problematisch anzusehen ist. Für die Dimensionen *Kompetenz* und *negative Einschätzung von Technikfolgen* aus dem Fragebogen zur Technikaffinität, wurden in diesem Versuch signifikante Unterschiede zwischen den Altersgruppen gefunden. Werden diese Ergebnisse mit den Ergebnissen aus Kapitel 4 verglichen, so fällt auf, dass bezüglich der Kompetenz insbesondere die älteren Frauen niedrige Wertungen abgeben haben. Der signifikante Unterschied zwischen den älteren und jüngeren Teilnehmern am Validierungsversuch in den Dimensionen der Technikaffinität kann somit aus der Unausgewogenheit der Geschlechter resultieren.

Für die Technikeinstellung fällt der große Unterschied in der Dimension *Skepsis* auf. Der Unterschied in der Dimension lässt sich allerdings durch die Unterschiede in der Art und Weise der Internetnutzung und durch Unterschiede in der Häufigkeit der Internetnutzung erklären. So nutzen nur 6 ältere Personen das Internet täglich, 1 ältere Frau nutzt das Internet 2 bis 3 Mal in der Woche, 1 andere ältere Frau etwa 1 Mal in der Woche, 1 ältere Frau etwa 1 Mal im Monat und 3 ältere Frauen gaben an, dass sie keinen Internetzugang haben. Bei den jüngeren Versuchspersonen nutzen alle täglich das Internet. Da die fehlende Möglichkeit das Internet zu nutzen, die Erfahrungen im Umgang mit dem Internet limitiert, können auch die Antwortmöglichkeiten in der Dimension *Skepsis* einschränken werden, weshalb die Ergebnisse für diese Dimension verzerrt sind.

Die Unterschiede in der Dimension *Kompetenz* (Technikeinstellung) sind mit großer Wahrscheinlichkeit abermals auf das Ungleichgewicht zwischen älteren Männern und Frauen zurückzuführen; die Erläuterung hierfür ist identisch mit der Erklärung für die Dimension *Kompetenz* (Technikaffinität). Für die Unterschiede in der Dimension *Angst* lässt sich keine eindeutige Erklärung finden. Czaja und Lee [45, S. 29] weisen darauf hin, dass ältere Menschen allgemein ängstlicher sind, wenn sie mit Computer interagieren müssen. Karavidas et al.

[125] fanden signifikante Unterschiede zwischen älteren Männern und Frauen, wobei ältere Frauen eine größere Ängstlichkeit berichteten. Demgegenüber wurde von Wild et al [251] berichtet, dass weder das Geschlecht noch das Alter der Grund für die höhere Ängstlichkeit ist, sondern der Bildungsgrad und das Vorwissen die Ängstlichkeit signifikant beeinflussen.

Für die Dimension *Interesse* ist anzunehmen, dass der Unterschied zwischen älteren und jüngeren Teilnehmern auf die Art der Rekrutierung zurückzuführen ist. So wurden die jüngeren Versuchspersonen des Validierungsversuchs und alle Teilnehmer des Versuchs in Kapitel 4 allein durch Aushänge bzw. E-Mails rekrutiert. Die älteren Teilnehmer des Validierungsversuches, wurden zum großen Teil über die persönliche Ansprache von bereits am Versuch teilgenommenen Probanden rekrutiert. Es ist zu vermuten, dass die Teilnehmer die durch die Aushänge angesprochen wurden, ein höheres eigenes Interesse an Technik haben, als dies die älteren Teilnehmer haben, die persönlich angesprochen wurden.

In der Dimension *Exploration* resultieren die Unterschiede zwischen den Altersgruppen nicht aus den Unterschieden im Geschlechterverhältnis. Wie in Abbildung 50 auf Seite 169 zu erkennen ist, existieren für diese Dimension keine Unterschiede zwischen den jüngeren Männern und Frauen. Nach Barnard et al. [10] wird Exploration von älteren Menschen zwar genutzt um die Bedienung eines Gerätes zu erlernen, allerdings verwenden Senioren häufig auch Handbücher oder fragen eine andere Person, um sich die Bedienung erklären zu lassen.

Wie schon im Kapitel 4 zeigen sich auch im Validierungsversuch signifikante Unterschiede im Interaktionsverhalten zwischen jüngeren und älteren Versuchspersonen. Die Unterschiede im Interaktionsverhalten betreffen die Berührungsdauer und den aufgewendeten Druck während der Interaktion.

Besonders interessant ist, dass die Interaktionsparameter aus dem Validierungsversuch für die Dimension *Druck* nicht signifikant von den Ergebnissen aus dem Versuch in Kapitel 4 abweichen und das, obwohl zwischen dem ersten Versuch und dem Validierungsversuch mehr als 4 Jahre liegen. Die älteren Versuchspersonen nutzen, wie schon im ersten Versuch, häufiger mehr Druck während der Interaktion mit dem Touchscreen. Dieser Umstand ist insofern bemerkenswert, da bei der Vorbefragung zum Thema *Funktionsweise eines Touchscreens* (Abschnitt 5.6.1.1) die Unterschiede zwischen den Altersgruppen nicht sehr ausgeprägt waren.

Werden die Unterschiede in der Dimension *Dauer einer Interaktion* betrachtet, so zeigt sich, dass die älteren Teilnehmer, wie schon im Versuch in Kapitel 4, signifikant häufiger längere Interaktionen nutzen. Zusätzlich zeigte sich, wie in Abschnitt 5.6.1.2 dargestellt, dass ältere Nutzer pro Interaktion eine deutlich höhere Distanz zwischen Aufsetzpunkt und Abhebe punkt haben, welche zu mehr Interaktionen führt,

die schwierig vom System bewertet werden können. Eine mögliche Erklärung für dieses Verhalten ist die Zunahme von altersbedingten Einschränkungen (siehe Abschnitt 1.3.2 und 3.1). So wird durch psychomotorische Einschränkungen die Ausführung einer Interaktion ungenauer und langsamer [27], was die beobachteten Abweichungen erklären kann.

5.6.2 *Simulation*

Die nachfolgend beschriebene Simulation dient der Validierung des Modells aus dem Kapitel 4 und der Überprüfung der ersten Hypothese. Um die Simulation nachvollziehbar und wiederholbar zu machen, werden in dem ersten Unterabschnitt die Modellierung des Systems, der Aufgabe und die Simulationsbedingungen beschrieben. Hieran schließt sich der Unterabschnitt an, der die Ergebnisse der Simulation darstellt und diese mit den empirischen Ergebnissen vergleicht. Den Abschluss des Abschnitts bildet das Zwischenfazit, in welchem die erzielten Ergebnisse diskutiert werden.

5.6.2.1 *Modellierung und Simulationsbedingungen*

Als Aufgabe für die Simulation wurde die in Abschnitt 5.4 beschriebene Teilaufgabe – das Anzeigen eines Fotos auf dem Fernseher – gewählt, weiterhin wird nur die Bedienung des Systems durch die XBMC-Remote-App simuliert.

Die Beschränkung auf diese Aufgabe und App resultiert aus den Beschränkungen der MeMo-Werkbank. So kann, in der aktuellen Implementierung das Nutzermodell keine Navigationsentscheidungen treffen, um in der auf dem Fernseher angezeigten Bedienoberfläche per Steuerkreuz zu navigieren, da ein Nutzermodell die Bedeutung von *rechts*, *links*, *hoch* oder *runter* nicht kennt. Es ist zwar möglich einen Navigationspfad (z. B.: *runter*, *rechts*, *rechts*, *bestätigen*) durch Aufteilung in Teilaufgaben (1. Teilaufgabe *runter*; 2. Teilaufgabe *rechts*, usw.) zu simulieren, allerdings ist dieses Vorgehen fehleranfällig und es widerspricht dem Sinn einer Simulation.

Für die Simulation wurde nur die XBMC-Remote-App in der Werkbank modelliert. Das Mediacenter und die Yatse-App wurden nicht modelliert, da eine sinnvolle Simulation aufgrund der genannten Beschränkungen der MeMo-Werkbank nicht erwartbar ist. Für die Simulation wurden nicht alle Oberflächen der XBMC-Remote-App modelliert. So wurde darauf verzichtet, die Menüs für das Einrichten von XBMC-Server zu modellieren. Zwar kann die Oberfläche zum Einrichten von XBMC relativ einfach, innerhalb von 3 Interaktionen, erreicht werden, allerdings wurde im empirischen Test durch den Versuchsleiter eingeschritten, sofern ein Teilnehmer das entsprechende Menü öffnete. Für die Simulation wurden insgesamt 15 Dialoge modelliert, welche in 20 verschiedenen Systemzustände verwendet

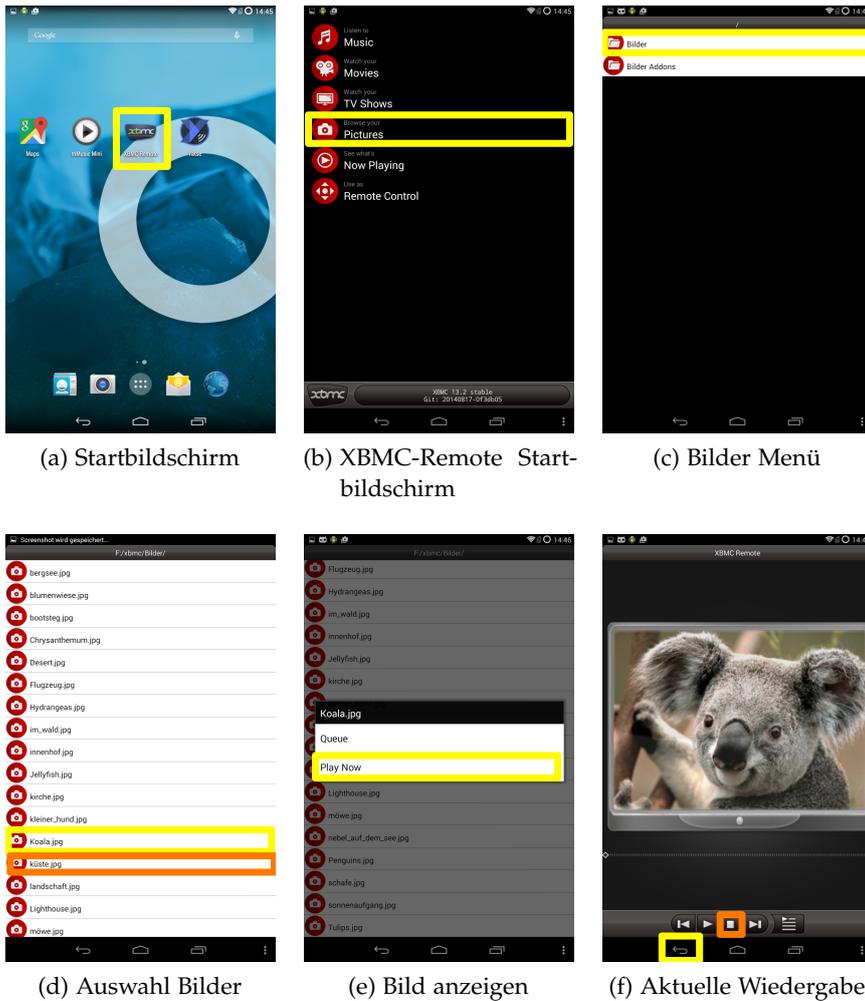


Abbildung 60: Oberfläche aus der verwendeten App

wurden. Die für die Simulation relevanten Dialoge und Systemzustände werden nachfolgend näher erläutert.

Wie im empirischen Versuch startet die Simulation auf dem Android-Hauptbildschirm (Abbildung 60a). Auf dem Startbildschirm muss der gelb umrahmte Button in der oberen Reihe kurz berührt werden, damit die App startet. Ist die App gestartet, wird der Startbildschirm der App (Abbildung 60b) angezeigt. Um auf dem Fernseher ein Bild anzeigen zu lassen, existieren zwei verschiedene Lösungswege. Der erste Lösungsweg nutzt das Steuerkreuz der App, welches unter dem Menüpunkt *Remote Control* zu finden ist (ohne Darstellung). Da die Simulation der Navigation mit Hilfe des Steuerkreuzes in der MeMo-Werkbank nicht möglich ist, wurde dieser Lösungsweg nicht modelliert. Für den zweiten Lösungsweg muss das Nutzermodell den Menüpunkt *Pictures* (gelb umrahmt in Abbildung 60b) aufrufen. Wird der entsprechende Button berührt, so öffnet sich eine neue Ansicht (Abbildung 60c), in welcher das Nutzermodell auf den Button *Bilder* tippen muss. Hiernach öffnet sich die Ansicht mit den verfügbaren Bildern (Abbildung 60d). In der Ansicht mit den Bildern muss das Nutzermodell das Bild mit der Bezeichnung *Koala* (Gelb gerahmt) suchen und durch längeres Berühren des Bildschirms ein Menü öffnen (Abbildung 60e), welches als Auswahloptionen das *Einfügen in eine Warteschlange* oder das *Abspielen* des Bildes anbietet. Jede der beiden Optionen führte dazu, dass das Bild auf dem Fernsehbildschirm angezeigt wird, und die Ansicht der App wechselt (Abbildung 60f). Um das nächste Bild anzeigen zu lassen, muss das Nutzermodell über den *Zurück*-Button (■) (Gelb umrandet) zur Auswahl der Bilder (Abbildung 60d) zurückkehren. In diesem View muss es das vorherige oder nächste Bild (Orange gerahmt) lange berühren, um es anzeigen zu lassen. Zum erfolgreichen Abschluss der Aufgabe soll auf dem Fernseher die Ansicht des Bildes geschlossen werden. Hierfür muss das Nutzermodell den *Stop*-Button (■) (Orange umrandet in Abbildung 60f) berühren.

Für die Simulation wurde die Aufgabenstellung in 5 Teilaufgaben, die sogenannten Subtasks, aufgespalten. Der erste Subtask (1) dient dazu, dass das Nutzermodell die XBMC-Remote-App startet. Die Anzeige der verfügbaren Bilder (Abbildung 60d) definierte das Ziel der zweiten Teilaufgabe (2). In der dritten Aufgabe (3) muss das Nutzermodell das Bild des Koalas auf dem Fernseher anzeigen lassen. Für die vierte Teilaufgabe (4) muss das Bild des Koalas durch das in der Liste vorherige oder nächste Bild ausgetauscht werden. Die fünfte und letzte Aufgabe (5) dient dazu, die Wiedergabe zu stoppen. Das für die Lösung der Aufgaben modellierte Aufgabenwissen ist in Tabelle 42 zusammengefasst. Die Spaltennamen entsprechen den Namen der Informationen (vgl. Kapitel 2). Die Werte in den Zeilen entsprechenden den initialen Ausprägungen der Informationen, wenn eine Teilaufgabe vom Nutzermodell begonnen wird. Zusätzlich zu den in

Subtask	Kategorie	Element	Aufgabe
1	-	XBMC remote	-
2	Bilder	Bilder	-
3	Bilder	Koala	Anzeigen
4	Bilder	Nächstes Vorheriges	Anzeigen
5	-	-	Schließen Beenden Zurück

Tabelle 42: Das Wissen des Nutzermodells für die 5 Subtasks der Simulation

der Tabelle aufgeführten Informationen ist für die Lösung der dritten Aufgabe (Koala anzeigen) eine weitere Information (*Hinweis = länger halten*) vorgesehen, die während der Simulation gesetzt wird, wenn das Nutzermodell den Button *Koala* mehrmals nur kurz berührt hat. Dieses Hinzufügen von Informationen entspricht dem Vorgehen im empirischen Versuch. Konnte eine Versuchsperson, nach mehrmaligem Versuchen, das Bild nicht auf dem Fernseher anzeigen lassen, so wurde der Hinweis gegeben, dass der Button länger zu halten war, um ein Optionsmenü (Abbildung 60e) anzeigen zu lassen. Damit das Nutzermodell den Hinweis anwenden kann, wurde im *Dialogdesigner* für den Button *Koala* und die Interaktion *Long tap* die Information *länger* annotiert. Für die ersten Versuche, das Bild anzeigen zu lassen, konnte das Nutzermodell die annotierte Information nicht nutzen, da keine entsprechende Information im Wissen des Nutzermodells hinterlegt war. Wurde nach mehreren erfolglosen Versuchen das entsprechende Nutzerwissen dynamisch hinzugefügt, so konnte das Nutzermodell die Interaktion *Long tap* mit der Lösung der Aufgabe in Verbindung setzen. Das gleiche Vorgehen wurde für das Schließen der Ansicht auf dem Fernseher gewählt (Hinweis: *Wiedergabe stoppen*) da hier ebenfalls im Versuch helfend eingegriffen werden musste.

Im nächsten Abschnitt werden die durch die Simulation erzeugten Interaktionsparameter vorgestellt und mit den empirischen Daten verglichen.

5.6.2.2 Interaktionsparameter

Für die Lösung der in Abschnitt 5.6.2.1 beschriebenen Aufgabe, sind mindestens 9 Interaktionsschritte notwendig. Die nachfolgende Auflistung beschreibt den optimalen Lösungspfad und die hierfür notwendigen Interaktionsschritte:

1. Starten der App (*Short tap*)
2. Öffnen Menüpunkt *Pictures* (*Short tap*)
3. Öffnen Menüpunkt *Bilder* (*Short tap*)
4. Aufruf Optionsmenü für das *Koala*-Bild (*Long tap*)
5. Auswahl der *Play Now*-Option (*Short tap*)
6. *Zurück*-Button auswählen (*Short tap*)
7. Aufruf Optionsmenü für das *Küsten* oder *Kleiner Hund*-Bild (*Long tap*)
8. Auswahl der *Play Now*-Option (*Short tap*)
9. *Stopp*-Button auswählen (*Short tap*)

Im empirischen Versuch und in der Simulation wurden der optimale Lösungspfad nicht beobachtet bzw. simuliert. Als besonders schwierig für die Versuchspersonen erwiesen sich das Aufrufen des Optionsmenüs, um ein Bild anzeigen zu lassen, der Wechseln zum nächsten bzw. vorherigem Bild und das Schließen der Anzeige auf dem Fernseher.

Da kein Versuchsteilnehmer das eingesetzte Mediacenter und die XBMC-Remote-App verwendet, wusste kein Teilnehmer, dass ein *Long tap* notwendig ist, um das Optionsmenü anzeigen zu lassen. Aufgrund des fehlenden Wissens führte keine der Versuchspersonen einen *Long tap* oder *Long press* als erste Aktion aus, um das Bild anzeigen zu lassen. Das Schließen der Ansicht auf dem Fernseher war aufgrund der Formulierung der Aufgabenstellung nicht einfach zu lösen. So stoppten nur wenige Probanden ($n = 2$) die Wiedergabe, um den Startbildschirm auf dem Fernseher anzeigen zu lassen, wohingegen ein Großteil der Probanden (62,5 %) den Zurück-Button auf dem Tablet verwendete, in der Annahme, dass hierdurch auch die Anzeige auf dem Fernseher ebenfalls verändert wird. Beim Anzeigen des vorherigen oder nächsten Bildes zeigte sich, dass es problematisch ist, wenn Buttons zur Navigation angezeigt werden, die keine Funktion besitzen.

In Abbildung 61 sind die benötigten Interaktionsschritte zur Lösung der Aufgabenstellung für die beiden Altersgruppen dargestellt. In Abbildung 61a sind die erfassten Interaktionen für den empirischen Versuch abgebildet. Wie der Abbildung zu entnehmen ist, haben die älteren Probanden ($n = 11$) durchschnittlich mehr Interaktionsschritte ($Mdn = 32$) benötigt als die jüngeren Probanden ($n = 6$; $Mdn = 28$), allerdings ist der Unterschied aufgrund der geringen Fallzahlen nicht signifikant ($U = 40,5$; $p > 0,05$). Die geringere Anzahl der in der Auswertung enthaltenen Versuchspersonen resultiert aus den Beschränkungen der Werkbank. So mussten Versuchspersonen aus der Auswertung

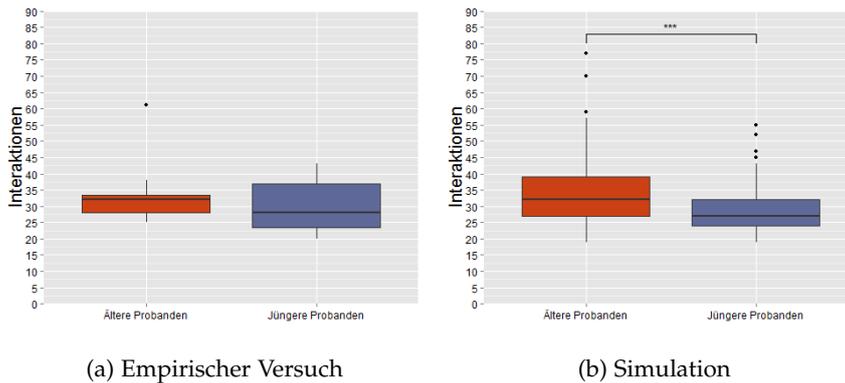


Abbildung 61: Anzahl der benötigten Interaktionen zur Lösung der Aufgabenstellung

ausgeschlossen werden, die das Steuerkreuz zur Lösung der Aufgabenstellung verwendeten, weil diese Interaktion von der Werkbank nicht simuliert werden kann.

In der Simulation, dargestellt in Abbildung 61b, zeigt sich ebenfalls ein Unterschied in der Anzahl der Interaktionsschritte. Wie im empirischen Versuch benötigten die älteren Nutzermodelle mehr Interaktionsschritte als die jüngeren Nutzermodelle. Für die Simulation ist der Unterschied hinsichtlich der benötigten Interaktionsschritte zwischen älteren Nutzermodellen ($n = 200$; $Mdn = 32$) und jüngeren Nutzermodellen ($n = 200$; $Mdn = 27$) hochsignifikant ($U = 28052$; $p < 0,001$), was auf die höhere Anzahl der Datenpunkte zurückzuführen ist. Ein Test der Effektstärke mittels Cliff's Delta [36] zeigt, dass der Unterschied ($|d| = 0,402$) mittelgroß ist (siehe Romano et al. [194]). Werden die empirischen Daten mit den Ergebnissen der Simulation verglichen, so zeigen sich weder für den Vergleich der jüngeren Nutzer mit den jüngeren Nutzermodellen ($U = 621$; $p > 0,05$), noch für die älteren Nutzer und älteren Nutzermodelle ($U = 1009$; $p > 0,05$), signifikante Unterschiede hinsichtlich der benötigten Interaktionsschritte.

Nachfolgend wird verglichen, wie die Versuchsteilnehmer und die Nutzermodelle mit dem System interagiert haben. Hierzu sei allerdings angemerkt: Das Betriebssystem des Tablets kann anhand von Messwerten (Dauer, Distanz zwischen Aufsetz- und Abhebeplatz, Größe der Berührungsfläche und Anzahl der Berührungspunkte) verschiedene Interaktionen wahrnehmen, allerdings kann das Betriebssystem nicht bestimmen, wie diese Werte von einer App verarbeitet werden. Die Interpretation einer Interaktion erfolgt durch die eingesetzte App selbst. Da durch das Betriebssystem nicht erfasst werden kann, wie die App die Eingabe auswertet, wurden die in den Abbildungen 62 und 63 dargestellten Häufigkeiten der Interaktionstypen wie nachfolgend beschrieben berechnet.

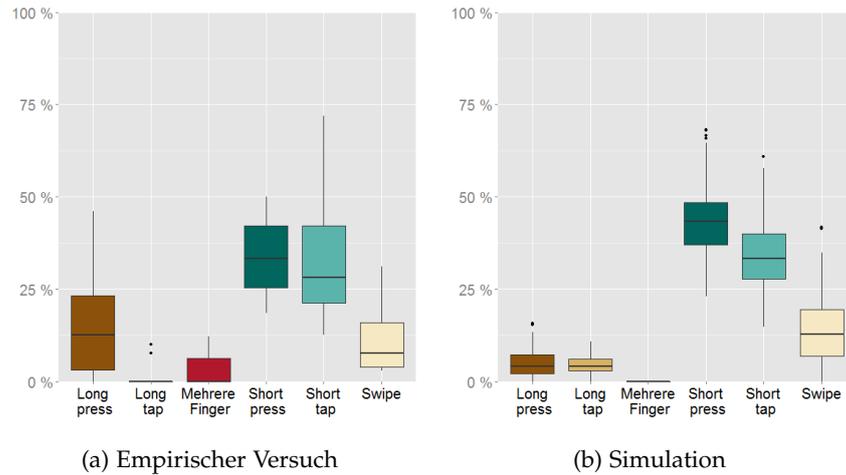


Abbildung 62: Häufigkeiten der Interaktionstypen für ältere Nutzer und Nutzermodelle

Als *Short tap* wurden Eingaben klassifiziert, die nicht länger als 250 ms dauerten, bei denen zwischen Aufsetz- und Abhebezeitpunkt weniger als 50 Pixel lagen und bei denen die erfasste Größe der Fingerfläche kleiner oder gleich der individuellen Barriere (siehe Tabelle 41) war. Als *Short press* wurden Interaktionen klassifiziert, die die Kriterien für den *Short tap* nur aufgrund der erfassten Fingergröße nicht erfüllten. Bei einem *Long tap* wurde der Touchscreen mit einem Finger länger als 250 ms berührt und die Position des Fingers änderte sich um nicht mehr als 50 Pixel. Weiterhin war die gemessene Fingerauflagefläche nicht größer als die in Tabelle 41 für die entsprechende Person definierte Barriere. Beim *Long press* verhält es sich wie beim *Short tap* und *Short press*, er weicht nur hinsichtlich des aufgebracht Drucks vom *Long tap* ab. Beim *Scrollen* berührte ein Finger den Touchscreen und die Positionsänderung betrug mehr als 50 Pixel. Eine Interaktion wurde als mit *mehreren Fingern* klassifiziert, wenn zeitgleich mehrere Berührungspunkte erfasst wurden.

In Abbildung 62 ist die Verteilung der Häufigkeiten der verwendeten Interaktionstypen für die älteren Versuchspersonen und Nutzermodelle dargestellt. Obwohl sich die beiden Grafiken ähneln, existieren für mehrere Interaktionstypen signifikante Unterschiede zwischen den Daten der Versuchspersonen und den Ergebnissen der Simulation. Um die Unterschiede zu zeigen, sind in Tabelle 43 die Häufigkeiten der Verwendung pro Interaktionstyp aufgeführt. Wie der Tabelle zu entnehmen ist, sind Interaktionen *Short press* und *Long tap* in der Simulation der älteren Nutzer überrepräsentiert und *Long press* unterrepräsentiert.

In Abbildung 63 sind die verwendeten Interaktionstypen für die jüngeren Versuchsteilnehmer bzw. Nutzermodelle dargestellt. Wie schon für die älteren Personen, so weisen auch die Darstellungen für

Interaktionstyp	Simulation	Versuch	Unterschied
Long press	M: 4,73 % SD: 3,82	M: 15,91 % SD: 16,19	$U = 1492$ $p = 0,046$
Long tap	M: 4,44 % SD: 2,52	M: 1,61 % SD: 3,62	$U = 455,5$ $p = 0,001$
Short press	M: 43,18 % SD: 9,02	M: 33,52 % SD: 10,06	$t(10) = -3,12$ $p = 0,009$
Short tap	M: 34,09 % SD: 9,19	M: 34,82 % SD: 19,48	$U = 973$ $p > 0,05$
Mehrere Finger	M: 0,0 % SD: 0,0	M: 3,11 % SD: 4,29	$U = 1600$ $p < 0,001$
Swipe	M: 16,56 % SD: 8,64	M: 11,04 % SD: 9,04	$U = 864$ $p > 0,05$

Tabelle 43: Häufigkeiten der verwendeten Interaktionstypen für ältere Nutzer und Nutzermodelle

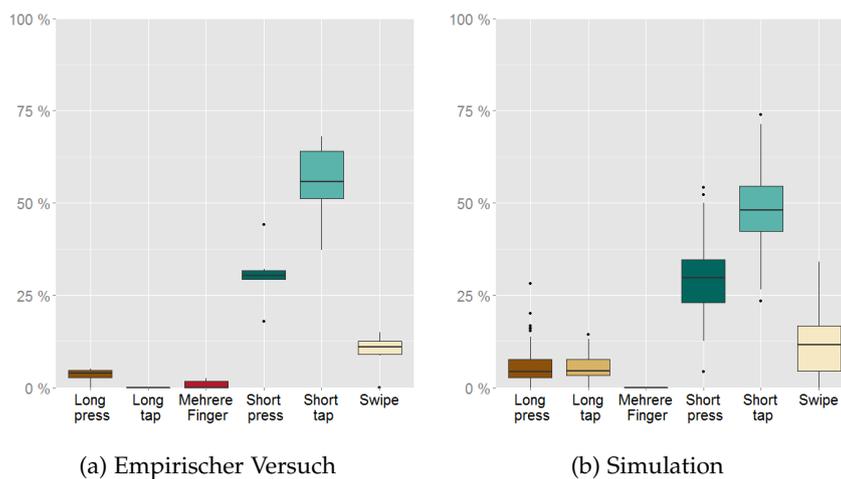


Abbildung 63: Häufigkeiten der Interaktionstypen für jüngere Nutzer und Nutzermodelle

Interaktionstyp	Simulation	Versuch	Unterschied
Long press	M: 5,14 % SD: 4,33	M: 3,30 % SD: 1,86	$U = 457$ $p > 0,05$
Long tap	M: 5,11 % SD: 3,03	M: 0,0 % SD: 0,0	$U = 75$ $p < 0,001$
Short press	M: 29,72 % SD: 8,86	M: 30,60 % SD: 8,36	$t(5) = 0,25$ $p > 0,05$
Short tap	M: 48,27 % SD: 9,59	M: 55,54 % SD: 11,38	$t(5) = 1,55$ $p > 0,05$
Mehrere Finder	M: 0,0 % SD: 0,0	M: 0,81 % SD: 1,26	$U = 800$ $p < 0,001$
Swipe	M: 11,76 % SD: 8,00	M: 9,75 % SD: 5,24	$U = 538$ $p > 0,05$

Tabelle 44: Häufigkeiten der verwendeten Interaktionstypen für jüngere Nutzer und Nutzermodelle

die jüngeren Personen große Ähnlichkeiten auf. Bei den statistischen Vergleichen zeigen sich, wie in Tabelle 44 dargestellt, signifikante Unterschiede bei einzelnen Interaktionstypen (*Mehrere Finger* und *Long tap*), wobei die am häufigsten genutzten Interaktionen in der Simulation sich nicht signifikant von denen im Versuch unterscheiden.

In einem weiteren Schritt dieses Vergleichs wird untersucht, ob die Simulation die im empirischen Versuch gefundenen Unterschiede, bezüglich der Verwendung von Interaktionstypen, reproduzieren kann.

Im empirischen Test existieren signifikante Unterschiede zwischen den Versuchspersonengruppen hinsichtlich der Verwendung von *Long press* ($t_{(10)} = -2,55$; $p = 0,028$) und *Short tap* ($t_{(14)} = 2,77$; $p = 0,014$). Die älteren Versuchspersonen nutzen signifikant häufiger *Long press* und die jüngeren Teilnehmer mehr *Short tap*-Interaktionen. In der Simulation existieren signifikante Unterschiede zwischen den Nutzermodellklassen für die Interaktionstypen *Long tap* ($U = 22893,5$; $p = 0,012$), *Short tap* ($U = 34165,5$; $p < 0,001$) und *Short press* ($t_{(397)} = -15,05$; $p < 0,001$). Die Simulation kann die in den empirischen Daten gefundenen Unterschiede somit nicht adäquat reproduzieren, weshalb nachfolgend untersucht wird, ob die Häufigkeit der Verwendung von Druck auf dem Touchscreen korrekt vorhergesagt werden kann, weil die Verwendung von Druck auf einem kapazitiven Touchscreen für ein unpassendes mentales Modell vom Eingabegerät spricht.

In Abbildung 64 ist dargestellt, wie häufig von den älteren Versuchsteilnehmern und älteren Nutzermodellen während der Lösung der

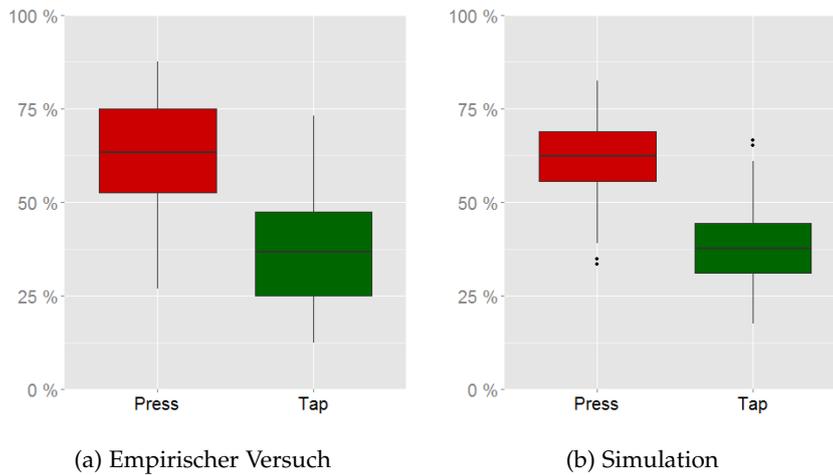


Abbildung 64: Häufigkeiten für Press- und Tap-Interaktionen bei den älteren Nutzern und Nutzermodellen

Bilderaufgabe die Interaktionen mit Druck ausgeführt wurden. Wie schon aus der Abbildung erkennbar, sind die Unterschiede hinsichtlich der Verwendung von Druck zwischen den älteren Versuchspersonen (Press: $M = 60,91$ %; $SD = 20,37$) und den älteren Nutzermodellen (Press: $M = 61,47$ %; $SD = 9,65$) nicht signifikant ($U = 1164,5$; $p > 0,05$).

Die Abbildung 65 stellt die Verwendungshäufigkeit von Press-Interaktionen für die jüngeren Versuchspersonen und jüngeren Nutzermodelle dar. Wie schon in Abbildung 64 sind sich die simulierten und empirisch gemessenen Ergebnisse sehr ähnlich. Bei den jüngeren Versuchspersonen wurde im Mittel in 44,03 % ($SD = 11,91$) der Fälle eine Interaktion mit Druck ausgeführt. In der Simulation wurde für die jüngeren Nutzermodelle im Mittel in 46,63 % ($SD = 10,41$) der Fälle die Interaktion druckvoll ausgeführt. Der Unterschied zwischen Simulation und empirischen Daten ist nicht signifikant ($t_{(5)} = -0,527$; $p > 0,05$).

Wird nun, wie bereits erwähnt, unterstellt, dass insbesondere die Verwendung von Druck auf einem kapazitiven Touchscreen ein Zeichen für ein unpassendes mentales Modell ist, so kann festgehalten werden, dass die Simulation die Fehler aufgrund von falschen mentalen Modellen des Eingabegerätes vorhersagen kann.

In der Tabelle 45 sind repräsentative Interaktionspfade von je zwei jüngeren und zwei älteren Teilnehmern aufgeführt, die in der Auswertung berücksichtigt wurden. Die Einträge sind wie folgt zu lesen. Der Text vor dem Schrägstrich (/) beschreibt das bediente Element. Wurde eine Wisch- oder Scroll-Geste ausgeführt, so ist als Element *Hintergrund* angegeben. Die Zahl nach dem Schrägstrich beschreibt die Häufigkeit. Die Beschreibung *Bilder/1* bedeutet, dass der Button *Bilder* (dargestellt in Abbildung 60c auf Seite 183) einmal berührt wur-

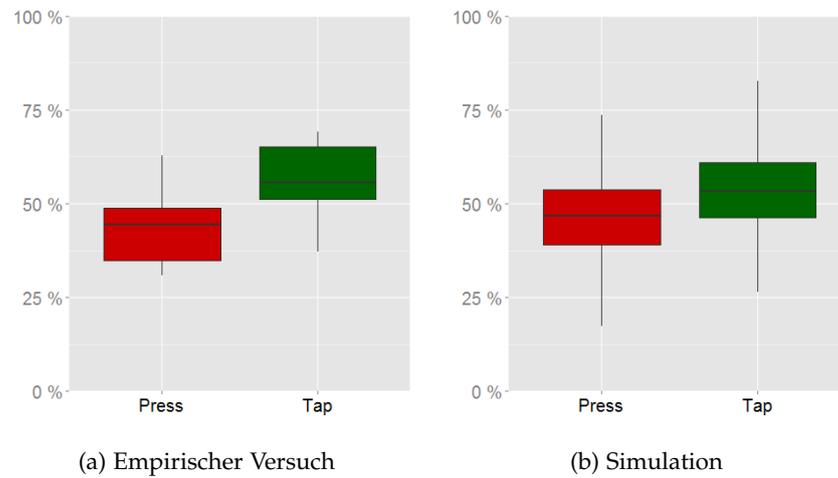


Abbildung 65: Häufigkeiten für Press- und Tap-Interaktionen bei den jüngeren Nutzern und Nutzermodellen

Gruppe	Interaktionssequenz
Jung	XBMC Remote/1, Pictures/1, Bilder/1, Hintergrund/1, Koala/2, Play Now/1, Hintergrund/1, Vorheriger/2, Hintergrund/1, Zurück/1, Kleiner Hund/1, Play Now/1, Zurück/3, Now Playing/1, Playlist/1, Zurück/1, Stopp/1
Jung	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Play Now/1, Nächster/1, Play/1, Playlist/1, Close/1, Zurück/2, Küste/2, Play Now/1, Zurück/3, Hintergrund/1, Remote Control/1, Zurück/1, Auswahl XBMC-Server/1, Hintergrund/1, Now Playing/1, Stopp/1
Alt	XBMC Remote/1, Pictures/1, Bilder/1, Hintergrund/1, Innenhof/1, Koala/5, Play Now/3, Vorheriger/4, Zurück/1, Kleiner Hund/5, Play Now/1, Stopp/1
Alt	XBMC Remote/1, Pictures/1, Bilder/1, Hintergrund/1, Koala/2, Play Now/1, Nächster/2, Play/1, Nächster/3, Playlist/1, Close/1, Zurück/1, Nächster/2, Vorheriger/1, Zurück/2, Küste/2, Play Now/1, Zurück/4, XBMC Remote/1, Remote Control/1, Stopp/1

Tabelle 45: Interaktionspfade von jüngeren und älteren Versuchspersonen

de, während *Koala/3* bedeutet, dass der *Koala*-Button (Abbildung 60d) dreimal hintereinander bedient wurde.

Um wichtige Interaktionsprobleme hervorzuheben, wurden diese in der Tabelle 45 farblich markiert (die Farbe impliziert keine Wertung). Die **gelben** Elemente kennzeichnen das Problem des *Long tap* zum Anzeigen von Bildern auf dem Fernseher. Wie am Anfang des Abschnitts beschrieben, musste zur Anzeige von Bildern auf dem Fernseher ein Button (*Koala*) länger gehalten werden. Wie in der Tabelle dargestellt, benötigten alle Teilnehmer im ersten Versuch mindestens zwei Interaktionen um das Optionsmenü anzeigen zu lassen, mit dessen Hilfe ein Bild auf dem Fernseher angezeigt werden kann. Mit **roter** Farbe hervorgehoben ist das Usability-Problem der nicht funktionierenden Navigationstasten. Nachdem das erste Bild auf dem Fernseher angezeigt wurde, waren die Versuchsteilnehmer angehalten auf das nächste bzw. vorherige Bild zu wechseln. Bei der XBMC-Remote funktionieren die Navigationstasten am unteren Rand (siehe Abbildung 60f) für Bilder nicht. Auch Wischgesten führen nicht zum gewünschten Ergebnis, allein der Wechsel zurück zum Auswahlbildschirm und das erneute Aufrufen des Optionsmenüs führen zum Erfolg. Mit **Blau** eingefärbt sind Schwierigkeiten beim Lösen der letzten Teilaufgabe (siehe Abschnitt 5.4); das Schließen der Bilderansicht auf dem Fernseher. Ein Großteil der Versuchspersonen nutzte den *Zurück*-Button des Tablets, um zum Startbildschirm der App zurückzukehren. Dort angekommen wurde von den Probanden bemerkt, dass auf dem Fernseher noch immer das letzte Bild angezeigt wurde. Nach dem Hinweis, dass zum Schließen der Ansicht die Wiedergabe gestoppt werden muss, konnte die Aufgabe von allen Teilnehmern gelöst werden.

In Tabelle 46 sind jeweils zwei Interaktionssequenzen für die älteren und jüngeren Nutzermodelle aufgeführt. Wie in der Tabelle 45, so lassen sich auch hier die im vorherigen Absatz angesprochenen Interaktionsprobleme finden. So nutzte kein Nutzermodell ein *Long tap* oder *Long press* als erste Interaktion beim *Koala*-Button. Ebenfalls kann in den Simulationsergebnissen beobachtet werden, dass das Problem der nicht funktionierenden Navigationstasten auch hier auftritt. Alle Nutzermodelle versuchen das nächste bzw. vorherige Bild über die entsprechenden Tasten bzw. Wischgesten aufzurufen. Auffallend ist, dass nach der *Zurück*-Interaktion es stellenweise beim Anzeigen des zweiten Bildes zur Wiederholung von Interaktionen kommt, wie diese auch in Tabelle 45 zu beobachten sind. Bei der unklaren Aufgabendefinition der letzten Teilaufgabe zeigen die Nutzermodelle ein Verhalten, welches den Versuchsteilnehmern ähnelt. Als Erstes werden verschiedene Button ausprobiert, da dies ohne Erfolg ist, wird der *Zurück*-Button des Tablets verwendet, um auf den Startbildschirm zurückzukehren. Auffällig bei den in Tabelle 46 aufgeführten Interaktionspfaden ist, dass alle Nutzermodelle die App zweimal über den

Nutzermodell	Interaktionssequenz
Jung	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Anzeigen/1, Hintergrund/1, Zurück/1, Kleiner Hund/1, Anzeigen/1, Zurück/4, XBMC Remote/1, Zurück/1, XBMC Remote/1, Now Playing/1, Stopp/1
Jung	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Anzeigen/1, Nächster/1, Vorheriger/1, Nächster/1, Vorheriger/1, Zurück/1, Kleiner Hund/1, Anzeigen/1, Hintergrund/3, Zurück/4, XBMC Remote/1, Zurück/1, XBMC Remote/1, Now Playing/1, Stopp/1
Alt	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Anzeigen/1, Nächster/1, Vorheriger/2, Hintergrund/2, Zurück/1, Küste/1, Anzeigen/2, Vorheriger/2, Nächster/1, Hintergrund/1, Vorheriger/1, Zurück/4, XBMC Remote/1, Zurück/1, XBMC Remote/1, Now Playing/1, Stopp/1
Alt	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Anzeigen/1, Nächster/1, Vorheriger/1, Hintergrund/2, Zurück/1, Hintergrund/1, Nächster/1, Hintergrund/4, Nächster/1, Zurück/1, Küste/2, Anzeigen/1, Hintergrund/1, Vorheriger/1, Hintergrund/1, Nächster/2, Hintergrund/1, Vorheriger/1, Hintergrund/2, Zurück/1, Hintergrund/1, Zurück/5, XBMC Remote/1, Zurück/2, XBMC Remote/1, Now Playing/1, Stopp/1

Tabelle 46: Interaktionspfade der jüngeren und älteren Nutzermodelle

Zurück-Button schließen, bevor sie die *Now Playing*-Anzeige aufrufen und die Wiedergabe stoppen. Dieses Verhalten ist der Modellierung der Aufgabe geschuldet. Erst nachdem das Nutzermodell die App zweimal geschlossen hatte, bekam es den Hinweis, dass die Ansicht mit dem Stopp-Button geschlossen werden kann.

5.6.2.3 Zwischenfazit

Bei der Simulation wurde darauf geachtet, dass das Systemmodell und das Aufgabenmodell möglichst realistisch abgebildet wurden. Durch Einschränkungen in der MeMo-Werkbank konnten nicht alle Interaktionsverläufe (Bedienung per Steuerkreuz) simuliert werden. Für die Interaktionsverläufe die simuliert werden konnten, wurde gezeigt, dass das Modell aus Kapitel 4 die im Validierungsversuch gemessenen Interaktionsparameter, für die Gruppe der jüngeren Personen bis auf kleine Abweichungen, reproduzieren kann. Für die Gruppe der älteren Versuchspersonen kann die Simulation zwar die Anzahl der Interaktionsschritte reproduzieren, allerdings existieren signifikante Unterschiede zwischen empirischen Daten und Simulationsergebnissen hinsichtlich der verwendeten Interaktionstypen (insbesondere für *Short press* und *Long press*). Die Abweichungen können durch verschiedene Faktoren erklärt werden. So konnte zum einen in Abschnitt 5.6.1.2 gezeigt werden, dass die langen und kurzen Interaktionen in den Altersgruppen zwischen dem Versuch in Kapitel 4 und dem Versuch in diesem Kapitel unterschiedlich sind, weshalb die Aufgaben im Validierungsversuch ein anderes Interaktionsverhalten verlangen. Zum anderen wurden die Daten für die Modellbildung mit Personen erhoben, die in der älteren Versuchspersonengruppe knapp unter 60 Jahre (Median 58 Jahre) alt waren (vgl. Abschnitt 4.1.3) und damit deutlich jünger als die älteren Teilnehmer im Validierungsversuch (Median: 72 Jahre).

Die in den Tabellen 45 und 46 dargestellten Interaktionspfade weisen eine hohe Übereinstimmung auf; weshalb die im Versuch für diese Aufgabenstellung gefundenen Usability-Probleme auch durch die Simulation entdeckt werden können. Werden die in Tabelle 46 aufgeführten Interaktionspfade kritisch betrachtet, so fallen Artefakte auf, die der Simulation und den gewählten Modellierungsentscheidungen angelastet werden können. So schließen die meisten Nutzermodelle die Aufgabe erst ab, nachdem sie die App zweimalig beendet haben und den modellierten Hinweis erhalten haben.

Bezüglich der ersten Hypothese kann festgehalten werden: Die im Validierungsversuch gemessenen Interaktionsparameter und beobachteten Interaktionspfade können mit der erweiterten MeMo-Werkbank durch Simulation erzeugt werden, allerdings existieren Abweichungen für die Gruppe der älteren Personen in den Verteilungen bzgl. der Verwendung von Interaktionstypen.

5.6.3 *Kombination der Ansätze*

In diesem Abschnitt werden die beiden Ansätze dieser Arbeit miteinander kombiniert. Durch die Verknüpfung des Bayes'schen Netzes mit den Modellen zur Vorhersage von Interaktionsfehlern soll die Qualität der Ergebnisse gesteigert werden. Da die Validierung der so erzielten Ergebnisse nicht durch einen empirischen Versuch erfolgen kann, werden die durch die Simulation erzeugten Interaktionsparameter nur hinsichtlich ihrer Konsistenz mit den zu erwartenden Ergebnissen bewertet.

Zur Simulation wurde das bereits für den Validierungsversuch entworfene Systemmodell der Android-App verwendet. Hierdurch entsprechen die Einschränkungen dieser Simulation den bereits in Abschnitt 5.6.2 beschriebenen Einschränkungen. Nachfolgend werden zuerst die erzeugten Nutzermodelle beschrieben, um im Anschluss die Interaktionsparameter darzustellen.

5.6.3.1 *Nutzermodelle*

Für die Simulation wurden zwei Nutzermodellgruppen ausgewählt. Für die jüngere Nutzermodellgruppe wurde das Alter auf 18 bis 29 eingegrenzt. Bei den älteren Nutzermodellen wurde nur das minimale Alter, mit 65 Jahren, definiert. Um die Nutzermodelle zu erzeugen, wurde das Bayes'sche Netz mittels einer Client-Server-Architektur (siehe Abschnitt 3.3.1) an die Werkbank angebunden. Für jede Nutzergruppe wurden 1000 Nutzermodelle generiert und mit diesen Modellen wurde die Interaktion simuliert. Für die Nutzermodelle werden die Alters- und Geschlechterverteilung in den Nutzermodellgruppen und die Verteilungen von Technikaffinität und Technikeinstellung nachfolgend dargestellt.

In der Abbildung 66 sind für die beiden Nutzermodellgruppen und für die demografischen Daten die Alters- und Geschlechterverteilungen dargestellt. Für die jüngeren Nutzermodelle (Abbildung 66a) ist das Verhältnis von Männern und Frauen über alle Altersjahre etwa gleich groß, eine Ausnahme bildet die Gruppe der 20-jährigen Nutzermodelle. Wird die durch das Netz generierte mit der demografischen Verteilung (Abbildung 66b) verglichen, so fallen kaum Unterschiede auf. Für die jüngeren Männer weichen die generierten Anteile von den demografischen Werten im Mittel um 0,55 % ab, für die jüngeren Frauen beträgt die mittlere Abweichung etwas weniger und liegt bei 0,26 %.

Bei den älteren Nutzermodellen (Abbildung 66c) ändert sich das Geschlechterverhältnis in Abhängigkeit vom Alter, wie es auch bei den demografischen Daten in Abbildung 66d zu beobachten ist. In beiden Verteilungen überwiegt der Anteil der weiblichen Personen/Modelle und hier insbesondere für die Personen/Nutzermodelle über 80 Jahren. Werden die Altersverteilungen verglichen, so fallen Unterschiede

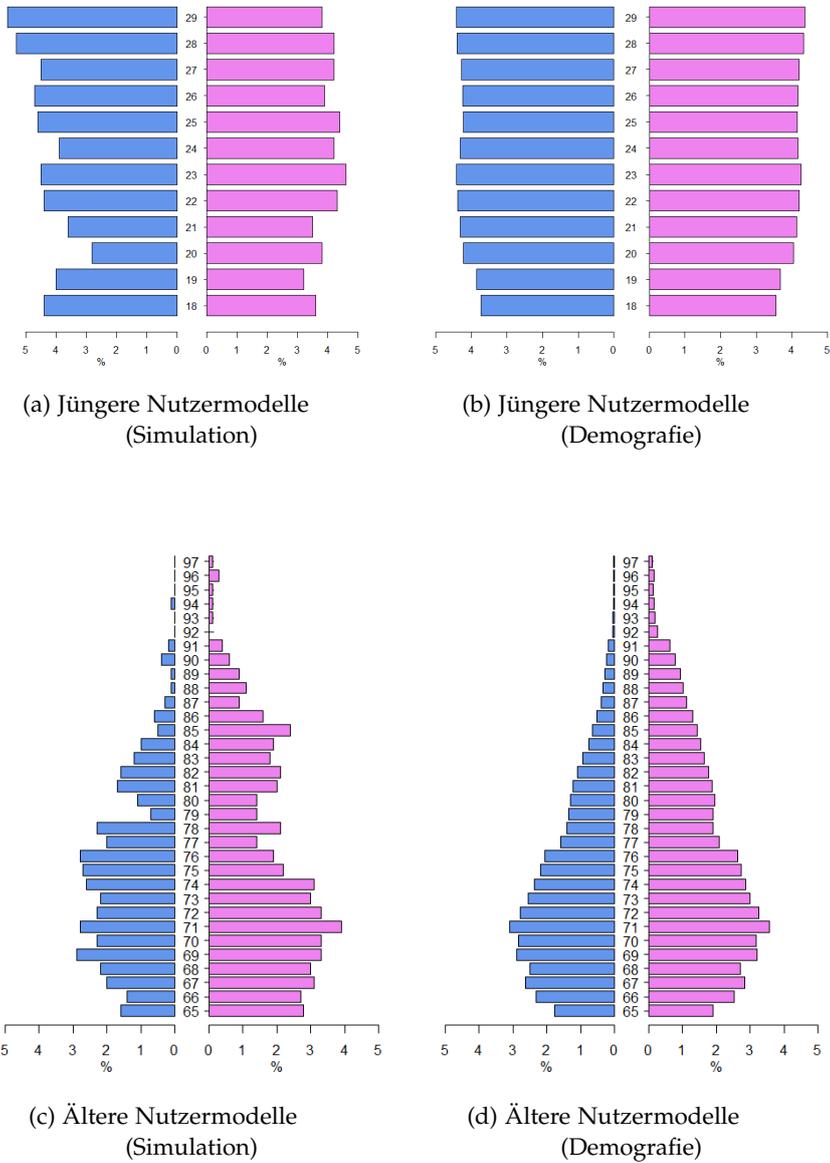


Abbildung 66: Alterspyramiden für die Nutzermodellgruppen und in der Bevölkerung. Die mit blau markierten Balken repräsentieren die männliche Bevölkerung, wohingegen die pinknen Balken für den Anteil der Frauen stehen.

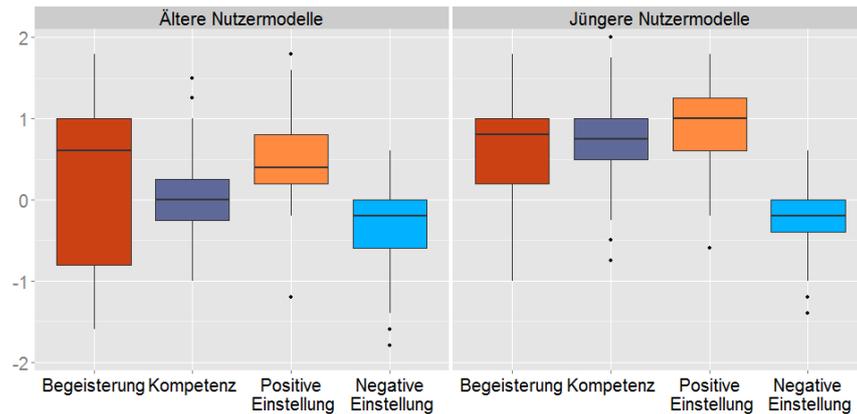


Abbildung 67: Verteilung für die Technikaffinität in den zwei Nutzermodellgruppen

zwischen den generierten und demografischen Daten für einzelnen Jahrgänge auf, allerdings sind die Unterschiede nicht sehr groß und betragen für die älteren Männer im Durchschnitt 0,31 % bzw. 0,28 % für die älteren Frauen.

Hinsichtlich der generierten Alters- und Geschlechterverteilungen ist anzumerken, dass sich diese weiter an die demografischen Verteilungen annähern, wenn mehr Simulationsdurchläufe durchgeführt und somit mehr Nutzermodelle generiert werden (siehe Abschnitt 3.3). Die in der Abbildung 66 auftretenden Abweichungen sind durch die eher geringe Anzahl von Iterationen (1000 pro Nutzermodellgruppe) geschuldet. Allerdings kann die Anzahl der Iterationen nicht beliebig gesteigert werden, da Rechenzeit und Speicherplatz die Anzahl der Iterationen begrenzen.

Nachdem das Alters- und Geschlechterverhältnis für die Simulation mit den demografischen Daten verglichen wurde, werden im nächsten Schritt die erzeugten Verteilungen für die Technikaffinität und Technikeinstellung betrachtet.

In Abbildung 67 sind die Verteilungen für die Dimensionen der Technikaffinität in den beiden Nutzermodellgruppen dargestellt. Ein Vergleich der zentralen Tendenz mittels Mann-Whitney-U-Test zeigt für jede Dimension hoch signifikante Unterschiede ($p < 0,001$) zwischen den Altersgruppen. Da aufgrund der Fallzahlen eine hohe Signifikanz schon bei kleineren Abweichungen zwischen den Nutzermodellgruppen auftreten kann, wurde zusätzlich die Effektstärke (Cliff's Delta) für jede Dimension berechnet. Für die Dimensionen *Negative Einstellung zur Technik* und *Begeisterung für Technik* ergeben sich zu vernachlässigen Unterschiede zwischen den Gruppen, da das Delta ($|d|$) kleiner als 0,147 ist (nach Romano et al. [194]). In der Dimension *Positive Einstellung zur Technik* ist der Unterschied zwischen den Nutzermodellgruppen ($|d| = 0,448$) mittelgroß und für die Dimension *Kompetenz im Umgang mit Technik* groß ($|d| = 0,684$). Dabei sind, wie

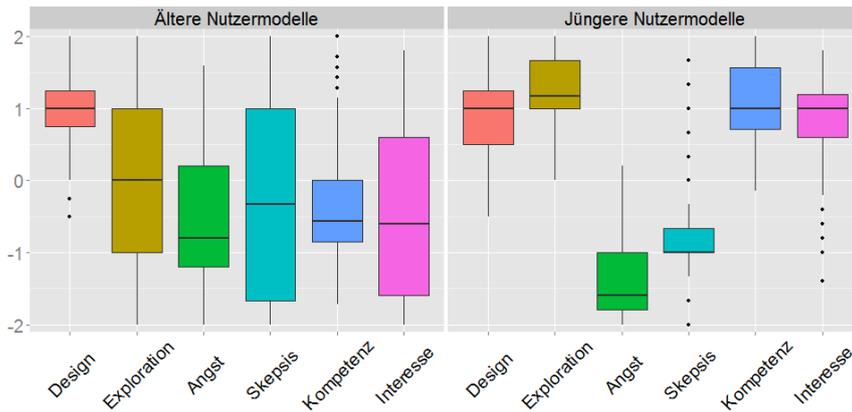


Abbildung 68: Verteilung für die Technikeinstellung in den zwei Nutzermodellgruppen

erwartet (siehe z. B. [127]), die Werte in der Dimension *Kompetenz* höher in der Gruppe der jüngeren Nutzermodelle. Ebenfalls nicht unerwartet sind die positiven Werte in der Dimension *Begeisterung im Umgang mit Technik* für beide Altersgruppen, da die Datenbasis im Rahmen von Experimenten erhoben wurde, an denen eher Personen teilnehmen, die an Technik interessiert sind. Bei der Dimension bezüglich der positiven Einstellung gegenüber der Technik zeigt sich, dass der Mittelwert bei den jüngeren Nutzermodellen im Durchschnitt höher als bei den älteren Nutzermodellen ist. Dieser Effekt ist ebenfalls nicht unerwartet und resultiert mit großer Wahrscheinlichkeit gleichfalls aus der Datenbasis.

Die Abbildung 68 zeigt die Verteilungen für die Dimensionen der Technikeinstellung in den Nutzermodellgruppen. Wie schon für die Technikaffinität sind die Unterschiede zwischen den Gruppen in allen Dimensionen hoch signifikant ($p < 0,001$). Bei der Berechnung von Cliff's Delta zeigen sich unterschiedliche Effektstärken in den einzelnen Dimensionen. Die Effektstärke und damit der Unterschied zwischen den Nutzermodellgruppen ist für die Dimensionen *Skepsis* ($|d| = 0,155$) und *Design* ($|d| = 0,173$) klein. Ein großer Unterschied existiert dafür in den Dimensionen *Kompetenz* ($|d| = 0,846$), *Exploration* ($|d| = 0,711$), *Ängstlichkeit im Umgang mit Technik* ($|d| = 0,616$) und *Interesse an ICT Technik* ($|d| = 0,567$).

Wie schon bei der Technikaffinität zeigen sich für die Technikeinstellung in einige Dimensionen die erwarteten Ausprägungen, so ist der Wert für die Kompetenz im Umgang mit Technik bei den älteren Nutzermodellen niedriger als bei den jüngeren Modellen. Für die Dimension *Exploration* zeigt sich ein bereits in der Literatur (vgl. [112, S. 55]) beschriebenes Verhältnis zwischen den Altersgruppen; so tendieren jüngere Personen eher zum Ausprobieren, während ältere Menschen häufiger die Bedienungsanleitung lesen bevor sie ein Gerät bedienen, um ein Verständnis von den Abläufen zu erhalten. Auch

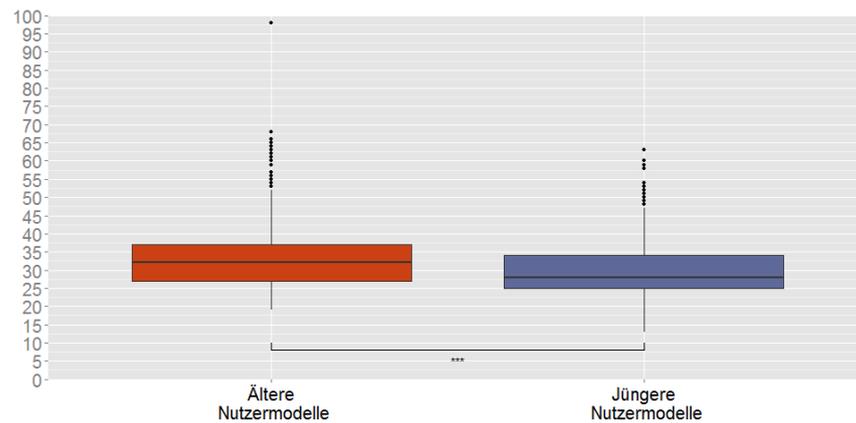


Abbildung 69: Verteilung für die Interaktionsschritte in den zwei Nutzermodellgruppen

in der Dimension *Angst* kann eine Verteilung beobachtet werden wie sie in der Literatur [44] beschrieben wurde. So sind die generierten Werte in der Gruppe der jüngeren Nutzermodelle deutlich niedriger als in der Gruppe der älteren Nutzermodelle. In der Dimension *Design* existieren nur minimale Unterschiede zwischen den älteren und jüngeren Nutzermodellen. In den Dimensionen *Skepsis* und *Interesse* fällt auf, dass die Varianz in den Werten für die älteren Nutzermodelle sehr groß ist. Der Grund für die große Varianz liegt in der Datenbasis für das Bayes'sche Netz begründet. Für die Modellierung der Nutzergruppe der Personen über 74 Jahre konnten zum einen nur wenige Daten genutzt werden und zum anderen hatten die Angaben der Versuchspersonen in den Fragebögen eine gegensätzliche Ausprägung.

In diesem Unterabschnitt wurden die Nutzerattribute betrachtet, wie sie von dem Bayes'schen Netz erzeugt und in der Simulation verwendet wurden. Wie gezeigt wurde, entsprechen die verwendeten Nutzerattribute den Erwartungen und bilden die demografisch Verteilungen gut ab.

5.6.3.2 Interaktionsparameter

In diesem Unterabschnitt werden die durch die Werkbank erzeugten Interaktionsparameter betrachtet. Als Erstes werden die Anzahl der Interaktionsschritte zwischen den Altersgruppen ausgewertet und es wird untersucht, welche Interaktionstypen während der Simulation von den Nutzermodellen verwendet wurden. Im zweiten Schritt wird ausgewertet wie die generierten Interaktionssequenzen aussehen und ob sie sich zwischen den Gruppen unterscheiden.

In Abbildungen 69 und 70 sind die Verteilungen für die Interaktionsschritte und Interaktionstypen in den zwei Nutzermodellgruppen dargestellt. Für die Anzahl der Interaktionsschritte kann ein signi-

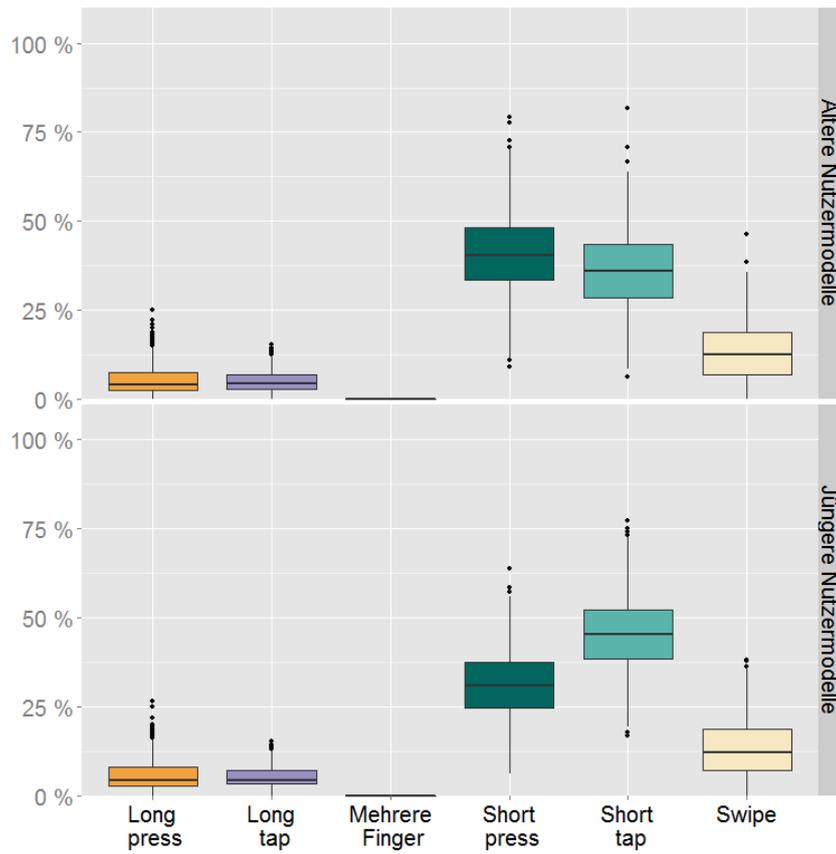


Abbildung 70: Verteilung für die Verwendung von Interaktionstypen in den zwei Nutzermodellgruppen

Interaktions- typ	Ältere Nutzer- modelle	Jüngere Nutzer- modelle	p	U	$ d $
Long press	M: 4,96 % SD: 4,16	M: 5,24 % SD: 4,32	0,046	474.394	0,051
Long tap	M: 4,75 % SD: 2,94	M: 4,86 % SD: 2,70	> 0,05	475.896	-
Short press	M: 40,82 % SD: 11,22	M: 31,2 % SD: 9,09	< 0,001	744.345	0,489
Short tap	M: 36,42 % SD: 10,72	M: 45,65 % SD: 10,40	< 0,001	268.824	0,462
Swipe	M: 13,05 % SD: 8,42	M: 13,03 % SD: 8,46	> 0,05	500.578,5	-

Tabelle 47: Verwendete Interaktionstypen der Nutzermodellgruppen

fikanter Unterschied ($U = 623711,5$; $p < 0,001$) zwischen der älteren (Mdn = 32) und jüngeren (Mdn = 28) Nutzermodellgruppen gefunden werden, allerdings ist der Unterschied zwischen den Gruppen nur klein ($|d| = 0,247$).

Für die Interaktionstypen sind in Tabelle 47 die Mittelwerte und Standardabweichungen zusammengefasst. Zusätzlich sind die p und U -Werte zur statistischen Signifikanz und Cliff's Delta ($|d|$) zur Bestimmung der Effektstärke in die Tabelle aufgenommen. Bei den *Long tap* und *Swipe*-Interaktionen besteht kein signifikanter Unterschied hinsichtlich der Verwendung zwischen den Gruppen. Bei den *Long press*-Interaktionen kann ein signifikanter Unterschied gefunden werden, allerdings ist das Delta zwischen den Verteilungen sehr klein, so dass der Unterschied vernachlässigt werden kann. Für die Verwendung von *Short press* und *Short tap* wurden hoch signifikante Unterschiede zwischen den Gruppen gefunden. Bei der Berechnung der Effektstärke ($|d|$) wurden für *Short press* große und für *Short tap* mittelgroße Unterschiede zwischen den Gruppen gefunden. Dieses Ergebnis entspricht den Erwartungen, da in beiden Versuchen (siehe Kapitel 4 und Abschnitt 5.6.1.2) festgestellt wurde, dass ältere Menschen mit mehr Druck auf einem Touchscreen interagieren.

Werden die erzielten Ergebnisse in dieser Simulation mit den Ergebnissen der Simulation ohne die Anbindung des Bayes'schen Netzes verglichen, so zeigt sich, dass hinsichtlich der Mittelwerte nur kleine Unterschiede existieren. Dieser Umstand ist nicht unerwartet, da die Verteilungen für Technikaffinität und Technikeinstellung im Validierungsversuch und als Ergebnis des Samplings ebenfalls ähnlich

sind und als Grundlage für die Berechnungen der Interaktionsarten ihre Verwendung finden (vgl. Abschnitt 4.2.4).

In Tabelle 48 sind die Interaktionssequenzen aufgeführt, die am häufigsten in der Simulation erzeugt wurden. Da für die älteren Nutzermodelle kein Interaktionspfad zweimal erzeugt wurde, erfolgte die Auswahl der in der Tabelle dargestellten Interaktionssequenzen zufällig. Wie bereits in Abschnitt 5.6.2.2 entsprechen die Markierung in der Tabelle der Klasse des durch die Werkbank gefundenen Usability-Problems, so sind die **gelb** markierten Elemente stellvertretend für das Problem des *Long tap* zum Anzeigen von Bildern auf dem Fernseher. Mit **roter** Farbe hervorgehoben ist das Problem der nicht funktionierenden Navigationstasten und **Blau** markiert die Schwierigkeiten, die durch die unklare Aufgabenstellung entstanden sind.

Zusätzlich zu den in Abschnitt 5.6.2.2 eingeführten Markierungen sind Interaktionselemente **fett** markiert, die besonders häufig wiederholt wurden oder zu einem unerwarteten Ergebnis führten. Die häufigen Wiederholungen beim ersten älteren Nutzermodell (markiert mit *) sind auf zwei verschiedene Probleme zurückzuführen. Zum einen wurden zuerst vom Nutzermodell nur die kurzen Interaktionen ausgeführt, die keine Aktion auslösen. Nach dem Hinweis, dass das Element lange gehalten werden muss, zeigte sich ein anderes Problem. Da das Nutzermodell (weiblich, 86, Tremor) starke Probleme hatte die Hand bzw. den Finger ruhig zu halten, wurden die *Long tap*-Interaktion häufiger als *Swipe* bzw. *Wischgeste* vom System interpretiert. Beim zweiten älteren Nutzermodell (weiblich, 85, Tremor, markiert mit †), wurden häufiger Buttons nicht getroffen. Beim Versuch den *Zurück*-Button zu bedienen wurde der *Home*-Button getroffen, weshalb der Startbildschirm von Android erschien. Auch für die Aufgabe die Wiedergabe zu stoppen brauchte das Nutzermodell drei Versuche, bevor der Button getroffen wurde und die Aufgabe abgeschlossen war.

Hinsichtlich der erzeugten Interaktionspfade kann festgehalten werden, dass diese die im empirischen Versuch gefundenen Usability-Probleme ebenfalls findet und zusätzlich noch auf Probleme hinweist, die auf körperliche Einschränkungen (Tremor) zurückzuführen sind.

5.6.3.3 Zwischenfazit

Wie in diesem Abschnitt gezeigt wurde, konnte die Güte der Vorhersage durch die Einbindung des Bayes'schen Netzes gesteigert werden. Die erzielten Ergebnisse der Simulation, insbesondere für die benötigten Interaktionsschritte, haben sich weiter den empirischen Daten angenähert. Es zeigt sich allerdings ebenfalls sehr deutlich, dass die in der App vorliegenden Usability-Probleme und die Probleme aus der unklaren Aufgabenstellung, die Unterschiede zwischen den Nutzergruppen überlagern und der Effekt der falschen oder fehlerhaften mentalen Modelle hierdurch verdeckt wird. Wie gezeigt wurde, ver-

Anzahl	Nutzermodell	Interaktionssequenz
4	Jung	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Anzeigen/1, Zurück/1, Küste/1, Anzeigen/1, Swipe/1, Zurück/4, XBMC Remote/1, Zurück/1, XBMC Remote/1, Now Playing/1, Stopp/1
3	Jung	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Anzeigen/1, Swipe/1, Zurück/1, Kleiner Hund/2, Anzeigen/1, Zurück/4, XBMC Remote/1, Zurück/1, XBMC Remote/1, Now Playing/1, Stopp/1
1	Alt	XBMC Remote/1, Pictures/1, Bilder/1, Koala/11*, Anzeigen/1, Swipe/1, Zurück/1, Kleiner Hund/1, Küste/3, Anzeigen/1, Swipe/1, Vorheriges Bild/1, Swipe/1, Zurück/5, XBMC Remote/2, Zurück/1, XBMC Remote/1, Now Playing/1, Stopp/1
1	Alt	XBMC Remote/1, Pictures/1, Bilder/1, Koala/3, Anzeigen/1, Swipe/2, Vorheriges Bild/1, Nächstes Bild/1, Zurück/1 [†] , XBMC Remote/1, Pictures/1, Bilder/1, Küste/1, Anzeigen/1, Zurück/1, Vorheriges Bild/1, Swipe/1, Nächstes Bild/1, Swipe/1, Zurück/5, XBMC Remote/1, Now Playing/1, Stopp/3 [†]

Tabelle 48: Interaktionspfade der jüngeren und älteren Nutzermodelle

wenden die älteren Nutzermodelle häufiger Interaktionen die mit Druck ausgeführt werden, was zu mehr Interaktionsfehlern führt.

Bezüglich der vom Bayes'schen Netz erzeugten Verteilungen kann, mit Verweis auf den Abschnitt 3.3.2, festgehalten werden, dass die Unterschiede in den Alters- und Geschlechterverteilungen allein auf die geringen Fallzahlen in der Simulation zurückzuführen sind und mit steigender Anzahl von Iterationen die Unterschiede zwischen empirischen Daten und generierten Daten verschwinden.

Zusammenfassend kann festgehalten werden. Die Kombination des Bayes'schen Netzes mit der erweiterten MeMo-Werkbank steigert die Qualität der Vorhersage und ergibt sinnvolle Interaktionsparameter. Somit kann auch die zweite Hypothese bestätigt werden.

5.7 FAZIT

In diesem Abschnitt der Arbeit wurde der Ansatz zur Simulation der mentalen Modelle durch einen unabhängigen Validierungsversuch verifiziert und es wurden das Bayes'sche Netz mit der Werkbank verbunden, wodurch die Vorhersagequalität weiter gesteigert werden konnte. Nachfolgend werden die positiven und negativen Aspekte aus diesem Kapitel zusammengefasst und bewertet.

5.7.1 Positive Aspekte

Der im Validierungsversuch eingesetzte Framework zum App-unabhängigen Aufzeichnen der Interaktion der Probanden mit dem Touchscreen liefert interessante Einblicke in das Interaktionsverhalten der unterschiedlichen Nutzergruppen und hilft die entwickelten Modelle zu validieren. Insbesondere die häufigen Interaktionen in der Gruppe der älteren Nutzermodelle bei denen knapp über 50 Pixel zwischen Aufsetz- und Abhebepunkt liegen, erklären viele Fehler in der Interaktion zwischen Mensch und System. Gleichfalls zeigt sich hierdurch, dass das in Abschnitt 4.2.5 vorgestellte Modell zur Manipulation der Benutzerschnittstelle auf richtigen Annahmen basiert und viele Fehler unbeabsichtigt passieren (Slips), die durch den Einsatz von entsprechenden Lösungen (z. B. Übertragung der Steadied-Bubble-Technologie [162] auf Touchscreens) reduziert werden können.

Ein weiteres Modell, welches durch den Validierungsversuch verifiziert werden konnte, ist das Modul zur *Verarbeitung der wahrgenommenen Ereignisse* (Abschnitt 4.2.4). Durch den in diesem Kapitel durchgeführten Versuch konnte gezeigt werden, dass die Simulation der mentalen Modelle von Eingabegeräten gültig ist. Ältere Menschen nutzen häufiger Druck, wenn sie mit einem Touchscreen einen Fernseher steuern müssen (vgl. Abschnitt 5.6.1.2), sie haben somit häufiger ein falsches Bild von der Funktionsweise eines kapazitiven Touchscreens. In der Simulation kann die häufige Nutzung von In-

teraktionen mit Druck und hieraus resultierende Fehler reproduziert werden (vgl. Abschnitt 5.6.2.2).

Der letzte Aspekt betrifft die Kopplung des Bayes'schen Netzes mit der MeMo-Werkbank. Wie in diesem Kapitel gezeigt werden konnte, liefert die Verknüpfung der beiden Konzepte sinnvolle Ergebnisse, die die Vorhersagen näher an die empirischen Ergebnisse bringen. Durch die Verwendung des Netzes in der Simulation wird es möglich die erzielten Ergebnisse in einen Kontext zu setzen und deren Auswirkungen bzw. Folgen zu quantifizieren, allerdings müssen hierfür noch weitere Arbeiten erfolgen.

5.7.2 Negative Aspekte

Bei den negativen Aspekten sind die Auswahl der teilnehmenden Probanden, die Einschränkungen in der Simulation, die Anforderungen an die Anzahl der notwendigen Iterationen und die noch offenen Erweiterungen der Werkbank zu nennen. Mit dem zuletzt genannten Punkt soll begonnen werden.

Um die Ergebnisse der Simulation quantifizieren zu können, müssen noch Erweiterungen an der MeMo-Werkbank vorgenommen werden. Zum einen muss die Werkbank um zusätzliche Regeln und Algorithmen erweitert werden, die die im Netz enthaltenen Informationen verwenden. Zum anderen muss die Werkbank dahingehend erweitert werden, dass mehr Iterationen pro Durchgang möglich sind. Die sinnvolle Verwendung des Bayes'schen Netzes in der Werkbank setzt voraus, dass sehr viele Iterationen (≥ 1.000) pro Nutzergruppe simuliert werden. Nur so ist es möglich, dass entsprechend viele Fälle (Einschränkungen) durch die Simulation abgedeckt werden. Eine entsprechend hohe Anzahl von Iterationen erfordert für die aktuelle Werkbank eine sehr leistungsfähige Hardware, um die Simulation durchzuführen und im Anschluss auswerten zu können.

Ein großes Problem in diesem Kapitel zeigte sich in der Simulation von indirekten Interaktionen. Wie in Abschnitt 5.6.2.1 dargestellt, kann die vorliegende Implementierung kein Steuerkreuz bedienen, da das Nutzermodell zwei Interaktionselemente in kein räumliches Verhältnis setzen kann. Aufgrund dieser Unzulänglichkeit konnten nicht alle Interaktionspfade simuliert werden, was die Anwendbarkeit der Ergebnisse einschränkt.

Ein weiteres Problem, welches insbesondere in diesem Kapitel von Bedeutung war, betrifft die Modellierung von Aufgabenwissen bzw. Hinweisen an das Nutzermodell während der Simulation. Im Validierungsversuch wurde den Versuchspersonen nach einer definierten Anzahl von Fehlversuchen ein Hinweis gegeben, wie das Problem gelöst werden kann. In der Werkbank muss dieser Hinweis durch eine zusätzliche Interaktion erfolgen, welche durch das Reporting gefiltert wird. Für zukünftige Implementierungen ist es wünschenswert, wenn

eine Möglichkeit implementiert und dokumentiert wird, die dieses Vorgehen standardisiert und vereinfacht.

Der letzte negative Punkt betrifft die Problematik der Probandenrekrutierung und hierbei sind insbesondere zwei Aspekte anzumerken. Zum einen haben mehr ältere Frauen als ältere Männer am Versuch teilgenommen, was zu schwer interpretierbaren Ergebnissen führt; zum anderen waren die Ausschreibungstexte für die Altersgruppen unterschiedlich, was ebenfalls zu Problemen führen kann.

Bezüglich des Ungleichgewichts zwischen den Geschlechtern bei den älteren Versuchspersonen ist festzuhalten: Wie bereits in Kapitel 1 beschrieben wurde ist das repräsentative Einwerben von Senioren sehr schwierig und zeitaufwendig, da meist nur die kognitiv und physisch fitten Personen teilnehmen. Im vorliegenden Versuch wurde über einen Monat nach passenden Versuchspersonen gesucht und trotz dieser intensiven Bemühungen war es nicht möglich mehr Männer zur Teilnahme am Versuch zu motivieren. Durch die stark unterschiedliche Anzahl von Männern und Frauen, können etwaige Unterschiede zwischen den Geschlechtern bei den älteren Teilnehmern nicht berücksichtigt werden.

Der Grund für die unterschiedlichen Ausschreibungstexte resultiert aus der Absicht, die Barrieren für die älteren Versuchspersonen möglichst niedrig zu halten, um auch Personen ohne Vorkenntnisse für die Teilnahme am Versuch zu gewinnen. Bei den jüngeren Versuchspersonen wurde davon ausgegangen, dass diese mit dem Touchscreen problemlos interagieren können. Diese Annahme wird durch die hohe Verbreitung von Mobiltelefonen und dem hohen Anteil von Smartphones in dieser Altersgruppe gestützt [246].

ZUSAMMENFASSUNG UND AUSBLICK

Das vorliegende Kapitel dient dazu, die umgesetzten Arbeiten im Überblick darzustellen und die erzielten Ergebnisse zu bewerten. Hierfür werden im nachfolgenden Abschnitt die erreichten Ziele zusammengefasst und mit den Zielsetzungen dieser Dissertation verglichen. Es werden weiterhin die Stärken und Schwächen der einzelnen Lösungen kritisch bewertet, um eine Beurteilung der Qualität zu ermöglichen. Das Kapitel schließt mit einem Ausblick über mögliche Erweiterungen und alternative Nutzungsmöglichkeiten der erzielten Ergebnisse.

6.1 ZUSAMMENFASSUNG

Das Ziel dieser Arbeit war es, einem Designer oder Ingenieur ein Werkzeug an die Hand zu geben, welches ihm ermöglicht die Usability von grafischen Benutzerschnittstellen für die Zielgruppe der älteren Personen zu testen, ohne Senioren zu einem Usability-Test einladen zu müssen (vgl. Kapitel 1).

Wie in Abschnitt 1.4.1 dargestellt existieren bereits verschiedene Werkzeuge (BS, DIAS und MeMo-Werkbank), die es einem Entwickler ermöglichen, die Auswirkungen von körperlichen Einschränkungen und fehlerhaften mentalen Modellen (bzgl. der Systemlogik) zu simulieren. Bei der Analyse der Werkzeuge zeigten sich allerdings offene Probleme, die die Anwendbarkeit einschränken.

So werden in zwei Werkzeugen (BS und DIAS) verschiedentlich Angaben zu Einschränkungen und deren Verteilungen in der Bevölkerung gemacht, allerdings sind die Angaben selten und nicht miteinander verknüpft. Hieraus resultieren zwei Probleme: So ist zum einen die Abschätzung der Relevanz eines Problems schwierig wenn keine demografischen Angaben über die Verteilung vorliegen, zum anderen treten verschiedene Einschränkungen im Alter häufig zusammen auf (Multimorbidität), weshalb die Abbildung der Zusammenhänge zwischen den Einschränkungen notwendig ist.

Hinsichtlich der mentalen Modelle konnte festgestellt werden, dass ein Werkzeug (MeMo-Werkbank) bereits dieses Konzept integriert, jedoch nur mentale Modelle bezüglich der Software (System Mental Models) abgedeckt werden. Mentale Modelle bezüglich der Eingabegeräte (Device Mental Models; DMM) werden von keinem der Werkzeuge in der Simulation verwendet. Eine Integration von DMMs in die Simulation ist notwendig, da unpassende oder fehlende DMMs das

Interaktionsverhalten beeinflussen und zu Abbrüchen der Interaktion führen können.

6.1.1 *Umgesetzte Arbeiten*

Für die Integration von demografischen Verteilungen in den Simulationsprozess wurde eine technische Lösung (Bayes'sches Netz; BN) entwickelt und mit Daten aus verschiedenen Quellen (Bücher, Studien und Versuchsergebnisse) initialisiert. Wie in Kapitel 3 dargestellt, werden verschiedene Arten von Einschränkungen (Wahrnehmung, Kognition, Motorik) und allgemeine Erkrankungen im Netz abgebildet und miteinander verknüpft, so dass Zusammenhänge zwischen Verteilungen erhalten bleiben. Durch die Programmierung einer Schnittstelle kann die MeMo-Werkbank auf das BN zugreifen und so demografische Daten in der Simulation verwenden.

Die Integration von DMMs in den Simulationsprozess wurde im Kapitel 4 beispielhaft für das Eingabegerät *Touchscreen* umgesetzt. Für die Integration der DMMs wurde dargestellt, wie die empirischen Daten erhoben wurden, wie die MeMo-Werkbank erweitert wurde und welche Ergebnisse mit der erweiterten MeMo-Werkbank erzielt werden können. Es wurden weiterhin verschiedene Module entwickelt, die den Wahrnehmungs-, Verarbeitungs- und Ausführungsprozess bei den Nutzermodellen während der Simulation beeinflussen. Zusätzlich wurde ein Modul entwickelt, das die vom Nutzermodell aufgebrachte Interaktion interpretiert und in eine durch das simulierte Eingabegerät auswertbare Interaktion transformiert.

6.1.2 *Erreichte Ziele*

Ein Ziel dieser Arbeit war es repräsentative Nutzermodelle zu erzeugen, die in der Simulation zur AUE verwendet werden können. Nach der Definition von Nutzermodellen in Abschnitt 1.4.2 auf Seite 18, müssen repräsentative Nutzermodelle verschiedenen Eigenschaften enthalten, die die Interaktion beeinflussen. Zusätzlich müssen Zusammenhänge zwischen den Eigenschaften abgebildet werden, um die Heterogenität von Nutzern in der gesamten Gesellschaft oder in einem Teil der Gesellschaft repräsentativ reproduzieren zu können.

Wie in Abschnitt 3.3.2 gezeigt ist das BN in der Lage das Wissen über die Eigenschaften von Personen und die Zusammenhänge zwischen den Eigenschaften zu speichern und Verteilungen korrekt wiedergeben zu können. Mit Hilfe des Bayes'schen Netzes und der Implementierung des Webservices können Daten über demografische Verteilungen in der Simulation genutzt werden und es kann vom Modellierer eine Abschätzung des Anteils der von einem Usability-Problem betroffenen Personen vorgenommen werden.

Die Integration von DMMs in die Simulation war das zweite Ziel dieser Arbeit. In den Kapiteln 4 und 5 wurde gezeigt, dass die Simulation der mentalen Modelle von Eingabegeräten möglich ist. Weiterhin konnte gezeigt werden, dass für die Integration von DMMs in die Simulation viele Veränderungen (Module zur Wahrnehmung, Verarbeitung und Ausführung) in der verwendeten Werkbank notwendig sind. Wie durch die simulierten Ergebnisse dargelegt werden konnte, kann die erweiterte MeMo-Werkbank fehlerhafte Interaktionen aufgrund von unpassenden oder fehlenden mentalen Modelle des Eingabegerätes nachahmen.

Das dritte Ziel der Arbeit war die Kombination der beiden Lösungen. Durch die Kombination der erweiterten MeMo-Werkbank mit dem Bayes'schen Netz können Einschränkungen von älteren Nutzern simuliert werden und aus den Ergebnissen der Simulation kann auf die Relevanz von Usability-Problemen geschlossen werden (vgl. Abschnitt 5.6.3).

6.1.3 Kritische Betrachtung der erreichten Ziele

In diesem Unterabschnitt werden die in dieser Arbeit erreichten Ziele hinsichtlich ihrer Stärken und Schwächen betrachtet. Weiterhin wird der Beitrag zur wissenschaftlichen Forschung erläutert.

6.1.3.1 Stärken

Wie in Kapitel 1 dargestellt fehlt es den vorhandenen Werkzeugen zur AUE mittels Simulation häufig an quantitativen Aussagen über Verteilungen von Einschränkungen in der Gesellschaft. Weiterhin werden unpassende bzw. fehlende mentale Modelle von Eingabegeräten und deren Auswirkungen auf die Interaktion bisher nicht in die Simulation integriert. Für beide Probleme konnten in dieser Arbeit entsprechende Lösungen gefunden werden.

Nach Persad et al. [178] wird umfangreiches Wissen über die späteren Nutzer und deren Einschränkungen benötigt, um Benutzerschnittstellen schon in frühen Phasen der Entwicklung evaluieren zu können. Insbesondere müssen Einschränkungen in der Wahrnehmung, Kognition und Motorik beachtet werden, um valide Aussagen treffen zu können. In verschiedenen Werkzeugen zur AUE (z. B. BS [17] und DIAS [78]) sind Daten über Einschränkungen hinterlegt, allerdings sind die Zusammenhänge zwischen den Einschränkungen nur unzureichend modelliert und die Verteilungen der betroffenen Personen in der Gesellschaft sind nicht für alle Einschränkungen hinterlegt.

Mit der in Kapitel 3 vorgestellten Lösung werden demografische Verteilungen über Einschränkungen in der Wahrnehmung, Kognition und Motorik korrekt wiedergegeben. Zusätzlich sind die Zusammenhänge zwischen den Einschränkungen bzw. Nutzerattributen modelliert, so dass Änderungen eines Attributes zu korrekten Veränderungen von

anderen Nutzerattributen führen. Durch diesen Umstand bietet das BN eine Flexibilität, die so durch andere Lösungen (BS, DIAS und HADRIAN) nicht erreicht werden kann.

Das zweite offene Problem, die Integration von DMMs in die Simulation, wurde in dieser Arbeit beispielhaft für das Eingabegerät *Touchscreen* gelöst. Wie von Hanson [87] beschrieben haben ältere Versuchspersonen häufiger unpassende oder fehlende Modelle von der Funktionsweise eines Eingabegerätes. Durch Einschränkungen in der Kognition fällt es ihnen schwerer neue Interaktionsformen zu erlernen [86], weshalb häufiger Interaktionsfehler beobachtet werden können, wenn für das Eingabegerät unpassende Interaktionsformen angewendet werden [68, 178]. In den bisherigen Werkzeugen zur AUE werden die DMMs nicht berücksichtigt (vgl. Abschnitt 1.4.1). In Kapitel 4 wurde gezeigt, welche Änderungen an der MeMo-Werkbank vorgenommen werden mussten, damit unpassende oder fehlende Modelle von Eingabegeräten simuliert werden können. Zusätzlich zu diesen Änderungen, die die Kognition betrafen, wurden Erweiterungen in der Wahrnehmung und Motorik vorgenommen, um die Auswirkungen von mangelhaftem Feedback und unbeabsichtigten Fehlern (Slips) simulieren zu können. Durch die Integration der DMMs und die zusätzlichen Erweiterungen konnten die im Versuch beobachteten Interaktionsparameter für ein Eingabegerät mit *Touchscreen* reproduziert werden. Weiterhin wurden für die älteren Nutzermodelle sehr unterschiedliche Interaktionspfade simuliert, die die Heterogenität in den Einschränkungen und die hieraus resultierenden Abweichungen in den Interaktionspfaden gut abdeckt. Wie in Kapitel 5 gezeigt werden konnte, sind die Erweiterungen auch für einen unabhängigen Versuch anwendbar und ermöglichen eine Abschätzung von Interaktionsfehlern aufgrund von unpassenden oder fehlenden DMMs.

Im Rahmen dieser Arbeit wurden nicht nur Lösungen für offene Probleme entwickelt, es wurden auch neue Erkenntnisse über das Interaktionsverhalten von jüngeren und älteren Versuchspersonen gefunden. Aus der Literatur ist bekannt, dass ältere Personen während der Interaktion den *Touchscreen* länger berühren (vgl. Stöfel et al. [225]) und das sie Probleme haben, wenn das Ziel klein ist und Druck während der Interaktion aufgewendet wird (vgl. Kobayashi et al. [139]). Die gleichen Ergebnisse konnten in den Versuchen in Kapitel 4 und 5 gefunden werden. Zusätzlich zu den bereits bekannten Erkenntnissen konnten gezeigt werden, dass die Interaktionen von älteren Personen auf einem *Touchscreen* häufig so ausfallen, dass die Interpretation durch das System schwierig ist. So konnte in Abschnitt 5.6.1.2 gezeigt werden, dass die Interaktionen von älteren Personen meist nur ein wenig länger als 250 ms dauern und dass oftmals nur ein wenig mehr als 50 px Abweichung zwischen Aufsetz- und Abhebe Punkt besteht, weshalb die Interaktion vom System häufig falsch erkannt wird.

6.1.3.2 Schwächen

Das in Kapitel 3 vorgestellte Bayes'sche Netz enthält in der aktuellen Ausprägung nur 111 Knoten. Mit dem Netz werden Einschränkungen in den Bereichen Wahrnehmung, Kognition und Motorik abgedeckt; zusätzlich sind Werte für Technikeinstellung und Technikaffinität gespeichert, allerdings enthält das Netz bisher nur Daten für Verteilungen aus Deutschland. Nach Kujala und Kauppingen [142] sollen neben den persönlichen Merkmalen auch aufgabenorientierte, geografische und soziale Eigenschaften einer Person eine Rolle spielen. Für das BN bedeutet dies, dass Informationen über Wissensstände in der Bevölkerung, anthropometrische Daten, Verteilungen in anderen Ländern und für weitere Nutzermerkmale, z. B. aus dem Bereich der Persönlichkeitspsychologie, aufgenommen werden müssen, um das Detaillevel der erzeugten Nutzermodelle weiter zu steigern.

Für die Integration von Device Mental Models in die Simulation wurde ein spezieller Fall (Steuerung eines Fernsehers via Touchscreen) statt einer allgemeinen Lösung ausgewählt. Die Beschränkung auf diesen Spezialfall erfolgte, weil einerseits im Versuch (siehe Kapitel 4) eben diese Interaktion sehr fehleranfällig war und die älteren Versuchspersonen hierbei Interaktionsmuster zeigten, die auf ein fehlendes oder unpassendes mentales Modell schließen lassen. Andererseits verhinderte das Fehlen von geeigneten abstrakten Modellen zur Beschreibung von Eingabegeräten eine allgemeinere Simulation. Zwar existieren verschiedene Klassifikationen, um Eingabegeräte systematisch zu beschreiben, allerdings sind diese Modelle zu alt (siehe z. B. Card et al. [29] und Jacob [110]) oder zu allgemein (siehe z. B. Dix [54]), um aktuelle Eingabegeräte adäquat und eindeutig beschreiben zu können. In 2011 wurde von Schulz et al. [209] versucht die Klassifikation von Jacobs durch Hinzufügen von weiteren Attributen zu aktualisieren, allerdings zeigte sich in der praktischen Anwendung, dass die Ergebnisse der Klassifikation eines Eingabegerätes durch verschiedene Experten nicht einheitlich sind, woraus der Schluss gezogen werden muss: Solange keine Klassifikation existiert, die einheitliche Ergebnisse liefert, können Fehler durch unpassende oder fehlende DMMs nicht allgemein beschrieben bzw. simuliert werden, weshalb auf Spezialfälle zurückgegriffen werden muss, die den Nachteil haben, dass deren Lebensdauer begrenzt ist.

Eine weitere Limitierung der Arbeit basiert auf der Auswahl der Versuchspersonen, die an den Versuchen teilnahmen. Wie sich in den Versuchen in Kapitel 4 und 5 zeigte, war es schwierig die geplanten Verteilungen von Personen umzusetzen. Für den Versuch in Kapitel 4 wurden nur wenige Versuchspersonen gewonnen, die älter als 65 Jahre waren. Im Versuch in Kapitel 5 konnten entsprechend mehr ältere Menschen zur Teilnahme motiviert werden, allerdings ist dort das Geschlechterverhältnis sehr unausgewogen. Beide Umstände führen dazu, dass die Ergebnisse schwer vergleichbar sind. Weiterhin muss

davon ausgegangen werden, dass die Repräsentativität hierdurch negativ beeinflusst wurde.

6.1.3.3 *Fazit*

In dieser Arbeit wurden zwei Beiträge vorgestellt und diskutiert, die den aktuellen Stand der Forschung erweitern. Zum einen wurde mit dem Bayes'schen Netz ein Werkzeug in den Bereich der AUE-Simulation eingeführt, welches ein hohes Maß an Flexibilität und Reliabilität aufweist. Durch die Anbindung des BNs an die MeMo-Werkbank können Angaben über demografische Verteilungen reproduziert und in der Simulation verwendet werden, was eine Abschätzung der Schwere eines Usability-Problems erlaubt.

Zum anderen wurde gezeigt wie Usability-Fehler mithilfe von DMMs simuliert werden können. Auch wenn die in Kapitel 4 Lösung nur einen Spezialfall abdeckt und hierdurch eine begrenzte Lebensdauer hat, so konnte dennoch gezeigt werden, dass eine Integration von DMMs sinnvoll ist und Erklärungen für Unterschiede in den Interaktionsparametern zwischen älteren und jüngeren Menschen liefern kann.

Es kann somit festgehalten werden: Die in dieser Arbeit vorgestellten Lösungen verbessern die Güte der Vorhersagen und sind geeignet, um frühzeitig Usability-Probleme durch Simulation zu finden. Es sind allerdings zusätzliche Arbeiten notwendig, um die Güte der Ergebnisse weiter zu steigern.

6.2 AUSBLICK

Die in dieser Arbeit vorgestellten Lösungen sind nicht nur für die Simulation in frühen Phasen der Entwicklung geeignet, es sind auch alternative Verwendungsmöglichkeiten denkbar. Darüber hinaus existieren offene Probleme, die weitere Arbeiten erfordern.

6.2.1 *Alternative Verwendungsmöglichkeiten*

Für die vorgestellte MeMo-Werkbank ist eine Erweiterung und Integration in eine Entwicklungsumgebung (IDE) denkbar. Wie in Kapitel 1 ausgeführt, wird die Simulation meist als zusätzliche Arbeit betrachtet, deren Mehrwert unterschätzt wird. Durch die Integration der Werkbank in eine IDE, können der Mehrwert gesteigert und Vorurteile abgebaut werden. Für DIAS existiert bereits eine Erweiterung, die es ermöglicht DIAS in der Netbeans-Entwicklungsumgebung zu verwenden [78], allerdings kann mit DIAS keine Abschätzung der Schwere eines Problems getroffen werden. Wenn die MeMo-Werkbank in eine IDE integriert wird, so muss der Entwicklungsprozess angepasst werden. Bisher werden die Datenstrukturen modelliert und daran

anschließend die Oberflächen zur Bedienung des Systems programmiert bzw. erzeugt. Wenn die Werkbank in eine IDE eingebunden wird, müssen zusätzlich die Oberflächen und logischen Verknüpfungen modelliert werden können. Der Vorteil der Integration ist, dass die für die Simulation aufgebrauchte Arbeit in der Implementierung des Systems gespart werden kann, da die modellierten Oberflächen und logischen Verknüpfungen bereits in Software überführt werden können. Ein ähnlicher Ansatz existiert bereits mit der Verbindung von MASP und MeMo-Werkbank (siehe Feuerstack et al. [66]), allerdings handelt es sich bei der MASP um eine modellbasierte Plattform zur Generierung von Benutzerschnittstellen und nicht um eine Entwicklungsumgebung.

Bezüglich des Bayes'schen Netzes sind ebenfalls verschiedene andere Verwendungsmöglichkeiten denkbar. Zum einen kann das Netz in der Analysephase verwendet werden, um relevante Einschränkungen der Zielgruppe zu evaluieren und Anforderungen an das Produkt zu definieren. Weiterhin kann das BN verwendet werden, um Personae realistischer zu formulieren. Eine realistische Beschreibung einer Person zu erarbeiten ist schwierig, insbesondere wenn keine Experten (z. B. Mediziner oder Gerontologen) im Team sind. So kann es leicht passieren, dass unrealistische Annahmen bzgl. möglicher Einschränkungen getroffen werden und aufgrund dieser Einschätzungen wichtige Produkteigenschaften nicht den realen Anforderungen entsprechen. Eine letzte Verwendungsmöglichkeit besteht in der Nutzung des Netzes zur Festlegung von Auswahlkriterien für Probanden. Ein BN kann einerseits verwendet werden, um eine Mindestanzahl von Probanden zu berechnen, wenn bestimmte Kriterien vorgegeben werden und die Stichprobe repräsentativ für die Gesellschaft sein soll. Zum anderen kann bei einer definierten Versuchspersonenzahl ein Sample erzeugt werden, welches für die gegebene Anzahl von Personen maximal repräsentativ ist.

6.2.2 Offene Arbeiten

Um den Nutzen der Kombination von Werkbank und Bayes'schem Netz zu maximieren, müssen weitere Algorithmen entworfen und Regeln zur Werkbank hinzugefügt werden. So ist es mit der aktuellen Werkbank nicht möglich die Auswirkungen von Einschränkungen in der Wahrnehmung oder Kognition adäquat zu simulieren, da keine entsprechenden Algorithmen bzw. Regeln existieren, die das Interaktionsverhalten beeinflussen.

Ein ähnlich gelagertes Problem besteht bei der Auswertung der durch die Simulation erzeugten Ergebnisse. In der vorliegenden Werkbank werden die Interaktionen durch Diagramme ähnlich den UML-Zustandsdiagrammen visualisiert. Eine Abweichung vom idealen Pfad wird farblich markiert, wobei keine Aussage über die Relevanz des

Problems getroffen wird. Zur Auswertung müssen deshalb die erzeugten Logfiles exportiert und in anderen Programmen (Excel, R) entsprechend ausgewertet werden. Es ist wünschenswert, wenn die Auswertungsmöglichkeiten in der Werkbank verbessert werden, um auf Probleme und deren Relevanz direkt hinzuweisen.

Um die Simulation von Fehlern aufgrund von unpassenden oder fehlenden mentalen Modellen der Eingabegeräte allgemeiner zu gestalten, muss eine allgemeine und valide Klassifikation entwickelt werden. Wie bereits erwähnt, sind die existierenden Klassifikationen nicht geeignet.

In diesem Teil des Anhangs werden die allgemeinen linearen Modelle zur Berechnung des mentalen Modells des Eingabegerätes dokumentiert. Die Performanz der Modelle ist in Tabelle 26 auf Seite 130 aufgeführt. In Tabelle 49 sind alle Abkürzungen alphabetisch geordnet, die in den allgemeinen linearen Modellen verwendet werden.

A.1 KAPAZITIVER TOUCHSCREEN

In diesem Abschnitt werden die beiden Formeln vorgestellt, die zur Berechnung der Wahrscheinlichkeit für einen kapazitiven Touchscreen verwendet werden. Die Formel 4 zeigt, wie die Wahrscheinlichkeit berechnet wird, wenn für die vorhergehende Interaktion ein Feedback wahrgenommen wurde. In der Formel 5 ist die Gleichung zur Berechnung der Wahrscheinlichkeit dargestellt, die verwendet wird, wenn kein Feedback wahrgenommen wurde.

$$\begin{aligned}
 P_{KT} = & 0,87760 + (Alt?|0;1) * (-0,26541 + BG_{TA} * 0,37295) + \\
 & Neg_{TA} * (-0,19824 + Pos_{TA} * 0,17827) + \\
 & Pos_{TA} * (0,42259 + Komp_{TA} * 0,24145 + A_{ICT} * 0,35559 + \\
 & Int_{ICT} * -0,21636) + BG_{TA} * (0,09866 + \\
 & Komp_{TA} * -0,06430 + A_{ICT} * -0,26446 + \\
 & Exp_{ICT} * -0,53592 + (Mann?|1;0) * -0,49623) + \\
 & Int_{ICT} * Komp_{TA} * 0,43874
 \end{aligned} \tag{4}$$

$$\begin{aligned}
 P_{KT-Fehler} = & 0,5607 + (Mann?|1;0) * (0,3676 + Komp_{TA} * -0,7180) + \\
 & Komp_{TA} * 0,4490 + A_{ICT} * -0,2598 + Int_{ICT} * \\
 & -0,7706 + Neg_{TA} * 0,3995 + (Alt?|0;1) * (A_{ICT} * \\
 & 0,3759 + Int_{ICT} * 1,1178 + Neg_{TA} * -0,7692)
 \end{aligned} \tag{5}$$

A.2 RESITIVER TOUCHSCREEN

Wie für den kapazitiven Touchscreen existieren auch für den resistiven Touchscreen zwei allgemeine lineare Modelle. Die Formel 6 wird verwendet, um die Wahrscheinlichkeit zu berechnen, wenn in der vorherigen Interaktion ein Feedback wahrgenommen wurde. Mit der

Abkürzung	Dimension
A_{ICT}	Technikeinstellung: Angst
BG_{TA}	Technikaffinität: Begeisterung
Exp_{ICT}	Technikeinstellung: Exploration
Int_{ICT}	Technikeinstellung: Interesse
$Komp_{TA}$	Technikaffinität: Kompetenz
Neg_{TA}	Technikaffinität: negative Einstellung zur Technik
Pos_{TA}	Technikaffinität: positive Einstellung zur Technik

Tabelle 49: Abkürzungen im allgemeinen linearen Modell

Formel 7 kann die Wahrscheinlichkeit berechnet werden, wenn kein Feedback wahrgenommen wurde.

$$\begin{aligned}
 P_{RT} = & 0,05778 + (Alt?|0;1) * 0,08958 + Komp_{TA} * 0,24446 + \\
 & Neg_{TA} * (0,33647 + Komp_{TA} * -0,12896 + A_{ICT} * 0,09840) + \\
 & BG_{TA} * (-0,27854 + Exp_{ICT} * 0,38631 + Int_{ICT} * 0,24618 + \\
 & Pos_{TA} * 0,13380) + Int_{ICT} * (Komp_{TA} * -0,50742 + \\
 & Exp_{ICT} * 0,08373) + Pos_{TA} * (Komp_{TA} * -0,27981 + \\
 & Neg_{TA} * -0,30002) + (Mann?|1;0) * (BG_{TA} * 0,16873)
 \end{aligned}
 \tag{6}$$

$$\begin{aligned}
 P_{RT-Fehler} = & 0,02493 + (Mann?|1;0) * -0,09306 + A_{ICT} * \\
 & (-0,24354 + Komp_{TA} * 0,74032) + Neg_{TA} * \\
 & (-0,05634 + Komp_{TA} * -1,66820 + Int_{ICT} * -0,36343) + \\
 & Int_{ICT} * (Komp_{TA} * -0,42536 + BG_{TA} * 0,16731) + \\
 & BG_{TA} * (Komp_{TA} * 0,84322 + Neg_{TA} * 0,61221) + \\
 & (Alt?|0;1) * (A_{ICT} * -0,32375 + Neg_{TA} * 0,62416)
 \end{aligned}
 \tag{7}$$

A.3 FERNBEDIENUNG

Die Formel 8 wird genutzt, um die Verwendungswahrscheinlichkeit für das mentale Modell einer Fernbedienung zu berechnen, wenn ein Feedback wahrgenommen wurde. Wenn kein Feedback wahrgenom-

men werden konnte, dann wird die Formel 9 zur Berechnung der Wahrscheinlichkeiten verwendet.

$$\begin{aligned}
 P_F = & 0.01119 + Neg_{TA} * (0,15616 + Komp_{TA} * 0,15228 + \\
 & A_{ICT} * 0,09284 + Int_{ICT} * 0,1404) + BG_{TA} * \\
 & (0,10359 + Komp_{TA} * 0,04027 + A_{ICT} * 0,12409 + \\
 & Neg_{TA} * -0,11607) + Int_{ICT} * Exp_{ICT} * 0,04177 + \quad (8) \\
 & Pos_{TA} * (A_{ICT} * -0,05476 + Exp_{ICT} * -0,03351 + \\
 & Neg_{TA} * -0,07952) + (Alt?|0;1) * (Neg_{TA} * -0,06799 + \\
 & BG_{TA} * -0,05951)
 \end{aligned}$$

$$\begin{aligned}
 P_{F-Fehler} = & 0,008694(Mann?|1;0) * (0,053272 + A_{ICT} * 0,024629 + \\
 & Int_{ICT} * -0,085268 + BG_{TA} * 0,095171) + \\
 & (Alt?|0;1) * -0,027233 + A_{ICT} * -0,012764 + \\
 & Int_{ICT} * (0,021139 + Komp_{TA} * 0,028906 + \\
 & A_{ICT} * -0,013017) + BG_{TA} * (0,017229 + \\
 & A_{ICT} * 0,064069) + Neg_{TA} * Int_{ICT} * 0,039492 \quad (9)
 \end{aligned}$$

BEISPIEL FÜR DEN EXPORT DER SIMULATIONSERGEBNISSE

In diesem Abschnitt des Anhangs werden Beispielausgaben des neuen Excel-Exporter-Moduls (EEM) dokumentiert. Das EEM exportiert die erzeugten Daten als Rohdaten, ähnlich einem Logfile, und in zusammengefasster Form.

B.1 ROHDATEN

In den Tabellen 50, 51 und 52 ist beispielhaft für einen Teil einer Simulation die Ausgabe der Rohdaten (Datenblatt Raw) dargestellt. Die Dokumentation des Arbeitsblattes musste in verschiedene Tabellen aufgeteilt werden, da sie sonst nicht darstellbar gewesen ist.

In Tabelle 50 sind die Spalten *ITERATION*, *STEP No.*, *State*, *UIObject Intended*, *Feedback*, *Probability UIObject*, *Device Mental Model* und *Interaction Intended* aufgeführt. Die Spalte *ITERATION* enthält die Identifikationsnummer des Nutzermodells. Die Identifikationsnummer ist pro Nutzergruppe und Simulation eindeutig; sie beginnt mit eins und wird pro Nutzermodell hochgezählt. In der Spalte *STEP No.* ist die Nummer des Interaktionsschrittes gespeichert. Die Spalte *State* enthält die Bezeichnung des modellierten Systemzustands, der dem Nutzermodell im entsprechenden Interaktionsschritt angezeigt wurde. In der Spalte *UIObject Intended* wird das vom Nutzermodell intendierte Interaktionsobjekt (Button, Link, usw.) gespeichert und die Spalte *Probability UIObject* enthält die Wahrscheinlichkeit, mit der das entsprechende Interaktionsobjekt gewählt wurde. Das im aktuellen Interaktionsschritt verwendete DMM wird in der Spalte *Device Mental Model* gespeichert. Abhängig vom DMM wird in der Spalte *Interaction Intended* die vom Nutzermodell intendierte Interaktion aufgeführt.

Die Fortsetzung der Tabelle 50 findet sich in Tabelle 51. In der Spalte *UIObject Executed* wird das Objekt gespeichert, welches während der Interaktion ausgeführt bzw. getroffen wurde. Da nicht nur die das intendierte Objekt vom ausgeführten Objekt abweichen kann, wird auch die ausgeführte Interaktion in der Spalte *Interaction Executed* gespeichert. Eine weitere Möglichkeit für einen Fehler ist, dass eine Interaktion falsch vom System verstanden wird. Diesem Umstand wird durch die Spalten *UIObject Understood* und *Interaction Understood* Rechnung getragen. In Abhängigkeit von der verstandenen Interaktion und dem getroffenen Objekt, wird vom System meistens ein Feedback gegeben, dessen Level in der Spalte *Feedback* gespeichert wird. Pas-

ITERATION	STEP No.	State	UIObject Intended	Probability UIObject	Device Mental Model	Interaction Intended	Probability Interaction
1	1	Start	XBMC-Remote	0,95	Resistive single press touchscreen	Long press	0,05
1	2	Start	XBMC-Remote	0,95	Resistive single press touchscreen	Short press	0,65
1	3	XBMC Start	Bilder	0,95	Capacitive multi-touch touchscreen	Short tap	0,45
1	4	XBMC Start	Bilder	0,95	Capacitive multi-touch touchscreen	Short tap	0,55
1	5	XBMC Bilder	Bilder	0,95	Resistive single press touchscreen	Short press	0,65
1	6	XBMC Bilder der Bilder	Bil- Koala	0,95	Resistive single press touchscreen	Short press	0,5
1	7	XBMC Bilder der Bilder	Bil- Koala	0,95	Capacitive multi-touch touchscreen	Short tap	0,45

Tabelle 50: Ausgabe der Rohdaten der Simulation

send dazu, wird in der Spalte *Feedback perceived* gespeichert, ob das Systemfeedback vom Nutzer wahrgenommen wurde.

Der letzte Teil der Ausgabe wird in Tabelle 52 dargestellt. Entspricht die Reaktion des Systems nicht der intendierten Interaktion, so wird in der Spalte *Error* markiert, dass ein Fehler aufgetreten ist. Die zweite Spalte mit der Bezeichnung *Error type* markiert die Art des Fehlers. Wurde vom Nutzermodell in der vorhergehenden Interaktion ein Fehler bemerkt, so kann es eine Aktion wiederholen, was in der Spalte *Redo action* vermerkt wird. In der Spalte *Necessary* wird vermerkt, ob eine Interaktion notwendig war, um eine Aufgabe zu lösen. Die Spalten *X intended*, *Y intended*, *X executed* und *Y executed* liefern die Koordinaten, wo das Nutzermodell plante den Bildschirm zu treffen (*intended*) und wo es den Bildschirm traf (*executed*).

B.2 ZUSAMMENGEFASSTEN DATEN

Der Export der zusammengefassten Daten dient dazu, dem Modellierer einen schnellen Überblick über die wichtigsten Nutzermodell- und Interaktionsparameter zu geben. Wie schon für das Tabellenblatt *Raw* musste auch die Übersicht für das Tabellenblatt *Userdata* in verschiedene Tabellen aufgebrochen werden.

In Tabelle 53 sind die ersten Spalten des Tabellenblatts dargestellt. In der ersten Spalte (*Iteration*) ist die Nummer des Nutzermodells gespeichert. In den Spalten *Age* und *Gender* sind das generierte Alter und das Geschlecht gespeichert. Zusätzlich wird in der Spalte *Usergroup* die Bezeichnung der Nutzergruppe aus der MeMo-Werkbank gespeichert. Die Spalten *TA_Enthusiasm*, *TA_Competence*, *TA_PositiveAttitude* und *TA_NegativeAttitude* enthalten die Mittelwerte für die 4 Dimensionen der Technikaffinität.

Die Tabelle 54 stellt alle Spalten dar, die die Daten für die Dimensionen zur Technikeinstellung enthalten.

In Tabelle 55 sind die Parameter enthalten, die etwas über die Interaktion zusammenfassen. Die erste Spalte (*Interactions*) speichert für jedes Nutzermodell die Anzahl der benötigten Schritte bis zur Lösung der Aufgabe oder bis zum Abbruch der Interaktion durch das Nutzermodell. In den Spalten *Slips* und *Mistakes* wird gespeichert, wie viele Interaktionsfehler in der Iteration aufgetreten sind, die auf einen Slip oder einen Mistake zurückzuführen sind. Hierauf aufbauend wird in den Spalten *OverallSlipRate* und *OverallMistakeRate* der Anteil der entsprechenden Fehler an den Gesamtschritten festgehalten. In der letzten Spalte (*OverallErrorRate*) wird die Gesamtfehlerrate gespeichert.

UIObject	Executed	Interaction	Executed	UIObject	Understood	Interaction	Understood	Feedback	Feedback perceived
XBMC-Remote		Long press		XBMC-Remote		Not understood		No Feedback	Wahr
XBMC-Remote		Short press		XBMC-Remote		Short tap		Normal	Wahr
Hintergrund		Short tap		Hintergrund		Short tap		No Feedback	Wahr
Bilder		Short press		Bilder		Short tap		Normal	Wahr
Koala		Short press		Koala		Short tap		Normal	Wahr
Koala		Short tap		Koala		Short tap		Low	Wahr
Koala		Short tap		Koala		Short tap		Low	Wahr

Tabelle 51: Ausgabe der Rohdaten der Simulation (Fortsetzung Tabelle 50)

Error	Error type	Redo action	Necessary	X intended	Y intended	X executed	Y executed
Wahr	Mistake	Falsch	Wahr	201	196	183	192
Falsch	-	Wahr	Wahr	201	196	182	212
Wahr	Slip	Falsch	Wahr	198	189	276	162
Falsch	-	Wahr	Wahr	198	189	266	196
Falsch	-	Falsch	Wahr	200	55	273	44
Falsch	-	Falsch	Wahr	200	460	159	449
Falsch	-	Falsch	Wahr	200	460	304	453

Tabelle 52: Ausgabe der Rohdaten der Simulation (Fortsetzung Tabelle 51)

Iteration	Age	Gender	Usergroup	TA_Enthusiasm	TA_Competence	TA_PositiveAttitude	TA_NegativeAttitude
1	26	Female	Young	0,2	-0,25	1,2	-1,4
2	27	Female	Young	0,6	0,75	0,6	-0,8
3	28	Male	Young	1,2	1	1,6	-1
4	18	Male	Young	0,6	0,5	0	0
5	27	Female	Young	0,6	0,75	0,6	-0,8
6	28	Female	Young	-1	-1	0,2	0
7	28	Female	Young	0	0,5	0,8	-0,8

Tabelle 53: Zusammengefasste Daten aus der Simulation

ICT_Surface	ICT_Exploratory	ICT_Anxiety	ICT_Scepticism	ICT_Competence	ICT_Interest
1,5	0,833	-0,8	-1,667	1,71	0,2
1,25	1,5	-1,4	-1,333	1,29	1
1	1,667	-2	-1	1,29	-0,2
1	0,333	0,2	1,333	0,71	0,2
1,25	1,5	-1,4	-1,333	1,29	1
0,25	-0,5	0	0,667	-0,43	-0,2
1	1	-1,8	-1,333	0,86	1,2

Tabelle 54: Zusammengefasste Daten aus der Simulation – Fortsetzung aus Tabelle 53

Interactions	Slips	Mistakes	OverallSlipRate	OverallMistakeRate	OverallErrorRate
20	2	1	0,1	0,05	0,15
25	0	0	0	0	0
22	1	2	0,05	0,09	0,14
13	0	0	0	0	0
18	1	0	0,06	0	0,06
26	2	1	0,08	0,04	0,12
19	1	0	0,05	0	0,05

Tabelle 55: Zusammengefasste Daten aus der Simulation – Fortsetzung aus Tabelle 54

LITERATUR

- [1] AGRAWAL, Y. ; PLATZ, E. A.; NIPARKO, J. K.: Prevalence of Hearing Loss and Differences by Demographic Characteristics Among US: Adults Data From the National Health and Nutrition Examination Survey, 1999-2004. In: *Archives of Internal Medicine* 168 (2008), Nr. 14, S. 1522–1530. – ISSN 0003–9926
- [2] AMT FÜR STATISTIK BERLIN-BRANDENBURG (HRSG.): *Statistischer Bericht: Schwerbehinderte Menschen im Land Brandenburg 2013*. Potsdam, April 2014
- [3] ANDERSON, J. R.; LEBIERE, C. : *The atomic components of thought*. Mahwah and N.J: Lawrence Erlbaum Associates, 1998. – ISBN 978–0–8058–2817–7
- [4] ANDERSON, J. R.; TAATGEN, N. A.; BYRNE, M. D.: Learning to achieve perfect timesharing: Architectural implications of Hazeltine, Teague, and Ivry (2002). In: *Journal of Experimental Psychology: Human Perception and Performance* 31 (2005), Nr. 4, S. 749–761. – ISSN 1939–1277
- [5] ANDROID OPEN SOURCE PROJECT: *Erläuterung zum Tool dumpsys*. <https://source.android.com/devices/input/dumpsys.html>. Version: 06.02.2015
- [6] ANDROID OPEN SOURCE PROJECT: *Quellcode ViewConfiguration.java*. https://github.com/android/platform_frameworks_base/blob/master/core/java/android/view/ViewConfiguration.java. Version: 13.11.2014
- [7] ANSORGE, U. ; LEDER, H. : *Wahrnehmung und Aufmerksamkeit*. 1. Wiesbaden: VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden GmbH, Wiesbaden, 2011 (Basiswissen Psychologie). – ISBN 978–3–531–16704–6
- [8] BAINBRIDGE, K. E.; HOFFMAN, H. J.; COWIE, C. C.: Diabetes and Hearing Impairment in the United States: Audiometric Evidence from the National Health and Nutrition Examination Survey, 1999 to 2004. In: *Annals of Internal Medicine* 149 (2008), Nr. 1, S. 1–10. – ISSN 0003–4819
- [9] BAKER, S. ; AU, F. ; DOBBIE, G. ; WARREN, I. : Automated Usability Testing Using HUI Analyzer. In: *19th Australian Conference on Software Engineering (aswec 2008)*, IEEE, 2008. – ISBN 978–0–7695–3100–7, S. 579–588

- [10] BARNARD, Y. ; BRADLEY, M. D.; HODGSON, F. ; LLOYD, A. D.: Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. In: *Computers in Human Behavior* 29 (2013), Nr. 4, S. 1715–1724
- [11] BARRETT, J. ; HERRIOTTS, P. : Running focus groups with older participants. In: McDONAGH, D. (Hrsg.); LANGFORD, J. D. (Hrsg.): *Focus groups*. London and New York: Taylor & Francis, 2003. – ISBN 978-0-415-26208-8, S. 115–144
- [12] BARRETT, J. ; KIRK, S. : Running focus groups with elderly and disabled elderly participants. In: *Applied Ergonomics* 31 (2000), Nr. 6, S. 621–629. – ISSN 00036870
- [13] BELLEBAUM, C. ; THOMA, P. ; DAUM, I. : Visuelle Wahrnehmung: Was, Wo und Wie. In: BELLEBAUM, C. (Hrsg.); THOMA, P. (Hrsg.); DAUM, I. (Hrsg.): *Neuropsychologie*. Wiesbaden: VS Verlag für Sozialwissenschaften, 2012. – ISBN 978-3-531-16827-2, S. 31–46
- [14] BICKEL, H. : Epidemiologie. In: WALLECH, C.-W. (Hrsg.); FÖRSTL, H. (Hrsg.): *Demenzen*. Stuttgart: Thieme, 2012 (Referenzreihe Neurologie Herausgegeben von G. Deuschl, H.C. Diener, H.C. Hopf). – ISBN 978-3-13-159492-1, S. 18–35
- [15] BIERMANN, H. ; WEISSMANTEL, H. ; INSTITUT FÜR ELEKTRO-MECHANISCHE KONSTRUKTIONEN (Hrsg.): *Regelkatalog SENSI-Geräte: Bedienungsfreundlich und barrierefrei durch das richtige Design*. <http://www.emk.tu-darmstadt.de/~weissmantel/sensi/sensi.html>. Version: 2003
- [16] BISDAS, S. ; DOST, P. ; IRO, H. ; ISSING, P. ; KECK, T. ; KINGMANN, C. ; KUMMER, P. ; LEUWER, R. ; LIMBERGER, A. ; DI MARTINO, E. ; MAUNE, S. ; MATTHEIS, S. ; PLINKERT, P. ; PLONTKE, S. ; REISS, M. ; REISS, G. ; SCHICK, B. ; SCHWAGER, K. ; SIEGERT, R. ; WALDFAHRER, F. ; VOGL, T. : Ohr. In: REISS, M. (Hrsg.): *Facharztwissen HNO-Heilkunde*. Heidelberg: Springer, 2009. – ISBN 978-3-540-89440-7, S. 139–298
- [17] BISWAS, P. ; ROBINSON, P. : Automatic evaluation of assistive interfaces. In: *Proceedings of the 13th international conference on Intelligent user interfaces*. New York and NY and USA: ACM, 2008 (IUI '08). – ISBN 978-1-59593-987-6, S. 247–256
- [18] BISWAS, P. ; ROBINSON, P. : Evaluating the design of inclusive interfaces by simulation. In: *Proceedings of the 15th international conference on Intelligent user interfaces*. New York and NY and USA: ACM, 2010 (IUI '10). – ISBN 978-1-60558-515-4, S. 277–280

- [19] BLUMENDORF, M. ; FEUERSTACK, S. ; ALBAYRAK, S. : Multimodal user interfaces for smart environments. In: LEVIALDI, S. (Hrsg.): *AVI '08 Proceedings of the working conference on Advanced visual interfaces*, 2008, S. 478–479
- [20] BONNEY, M. C.; CASE, K. ; HUGHES, B. J.; KENNEDY, D. N.; WILLIAMS, R. W.: *Using SAMMIE for computer-aided workplace and work task design*. 1974
- [21] BRAVE, S. ; NASS, C. : Emotion in Human-Computer Interaction. In: JACKO, J. A. (Hrsg.); SEARS, A. (Hrsg.): *The human-computer interaction handbook*. Mahwah and N.J: Lawrence Erlbaum Associates, 2003 (Human factors and ergonomics). – ISBN 978-1-4106-0672-3, S. 81–96
- [22] BRUNNER, C. : *Der Einfluss von anthropometrischen Parametern und der Körperzusammensetzung auf den Knochenstatus älterer Menschen unter Berücksichtigung des Alters und von Lebensstilfaktoren: Eine Untersuchung im Rahmen der Gießener Senioren Langzeitstudie*. Gießen, Justus-Liebig-Universität, Diss., 2012
- [23] BUGG, J. M.; ZOOK, N. A.; DELOSH, E. L.; DAVALOS, D. B.; DAVIS, H. P.: Age differences in fluid intelligence: Contributions of general slowing and frontal decline. In: *Brain and Cognition* 62 (2006), Nr. 1, S. 9–16. – ISSN 02782626
- [24] BÜLOW, J. : *Grundwissen Übergewicht und Adipositas: Folgen, Ursachen, Therapie und Fallstudie zu Ernährungs- und Bewegungsangeboten an Schulen*. Hamburg: disserta Verlag, 2012. – ISBN 978-3-9542502-0-2
- [25] BUSCH, M. ; MASKE, U. ; RYL, L. ; SCHLACK, R. ; HAPKE, U. : Prävalenz von depressiver Symptomatik und diagnostizierter Depression bei Erwachsenen in Deutschland. In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 56 (2013), Nr. 5-6, S. 733–739. – ISSN 1436-9990
- [26] BUSCH, M. ; SCHIENKIEWITZ, A. ; NOWOSSADECK, E. ; GÖSSWALD, A. : Prävalenz des Schlaganfalls bei Erwachsenen im Alter von 40 bis 79 Jahren in Deutschland. In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 56 (2013), Nr. 5-6, S. 656–660. – ISSN 1436-9990
- [27] CAPRANI, N. ; O'CONNOR NOEL E.; GURRIN CATHAL: Touch screens for the older user. In: AUAT CHEEIN, F. A. (Hrsg.): *Assistive Technologies*. INTECH Open Access Publisher, 2012, S. 95–118
- [28] CARD, S. K.: The model human processor: A model for making engineering calculations of human performance. In: *Procee-*

dings of the Human Factors and Ergonomics Society Annual Meeting
Bd. 25, 1981, S. 301–305

- [29] CARD, S. K.; MACKINLAY, J. D.; ROBERTSON, G. G.: The design space of input devices. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* ACM, 1990, S. 117–124
- [30] CARD, S. K.; MORAN, T. P.; NEWELL, A. : Computer text-editing: An information-processing analysis of a routine cognitive skill. In: *Cognitive Psychology* 12 (1980), Nr. 1, S. 32–74. – ISSN 00100285
- [31] CARD, S. K.; MORAN, T. P.; NEWELL, A. : The keystroke-level model for user performance time with interactive systems. In: *Communications of the ACM* 23 (1980), Nr. 7, S. 396–410. – ISSN 00010782
- [32] CARD, S. K.; MORAN, T. P.; NEWELL, A. : *The psychology of human-computer interaction*. Hillsdale and N.J.: L. Erlbaum Associates, 1983. – ISBN 978-0-89859-859-9
- [33] CEBALLOS-BAUMANN, A. O.: Idiopathisches Parkinson-Syndrom: Grundlagen, Medikamente, Therapieeinleitung. In: CEBALLOS-BAUMANN, A. O. (Hrsg.); ABELE, M. (Hrsg.): *Bewegungsstörungen*. Stuttgart: Georg Thieme Verlag, 2005. – ISBN 978-3-13-102392-6, S. 33–70
- [34] CHARNIAK, E. : Bayesian Networks Without Tears: Making Bayesian Networks More Accessible to the Probabilistically Un-sophisticated. In: *AI Magazine* 12 (1991), Nr. 4, S. 50–63
- [35] CHOU, R. ; DANA, T. ; BOUGATSOS, C. ; FLEMING, C. ; BEIL, T. ; AGENCY FOR HEALTHCARE RESEARCH AND QUALITY (Hrsg.): *Screening adults aged 50 years or older for hearing loss: A Review of the Evidence for the US Preventive Services Task Force: Evidence Synthesis No. 83. AHRQ Publication No. 11-05153-EF-1*. Rockville and MD, März 2011
- [36] CLIFF, N. : *Ordinal methods for behavioral data analysis*. Erlbaum, 1996. – ISBN 978-0-8058-1333-3
- [37] COHEN, J. : A Coefficient of Agreement for Nominal Scales. In: *Educational and Psychological Measurement* 20 (1960), Nr. 1, S. 37–46
- [38] CONGDON, N. G.; O'COLMAIN, B. J.; KLAVER, C. C. W.; KLEIN, R. ; MUÑOZ, B. ; FRIEDMAN, D. S.; KEMPEN, J. H.; TAYLOR, H. R.; MITCHELL, P. ; HYMAN, L. : Causes and Prevalence of Visual Impairment Among Adults in the United States. In: *Archives of Ophthalmology* 122 (2004), Nr. 4, S. 477–485. – ISSN 0003-9950

- [39] CONGDON, N. G.; VINGERLING, J. R.; KLEIN, B. E. K.; WEST, S. K.; FRIEDMAN, D. S.; KEMPEN, J. H.; O'COLMAIN, B. J.; WU, S.-Y. ; TAYLOR, H. R.; WANG, J. J.: Prevalence of Cataract and Pseudophakia/Aphakia Among Adults in the United States: The Eye Diseases Prevalence Research Group. In: *Archives of Ophthalmology* 122 (2004), Nr. 4, S. 487–494. – ISSN 0003–9950
- [40] CONRAD, N. : *Lebensziele in der neuropsychologischen Rehabilitation: Empirische Befunde und therapeutische Implikationen*. Marburg, Philipps-Universität, Diss., 2010
- [41] COOPER, A. : *The inmates are running the asylum*. Indianapolis and IN: Sams, 2004. – ISBN 978–0–672–32614–1
- [42] CRUICKSHANKS, K. J.; KLEIN, R. ; KLEIN, B. E. K.; WILEY, T. L.; NONDAHL, D. M.; TWEED, T. S.: Cigarette Smoking and Hearing Loss. In: *JAMA* 279 (1998), Nr. 21, S. 1715–1719. – ISSN 0098–7484
- [43] CRUICKSHANKS, K. J.; WILEY, T. L.; TWEED, T. S.; KLEIN, B. E. K.; KLEIN, R. ; MARES-PERLMAN, J. A.; NONDAHL, D. M.: Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin the epidemiology of hearing loss study. In: *American Journal of Epidemiology* 148 (1998), Nr. 9, S. 879–886
- [44] CZAJA, S. J.; CHARNES, N. ; FISK, A. D.; HERTZOG, C. ; NAIR, S. N.; ROGERS, W. A.; SHARIT, J. : Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). In: *Psychology and Aging* 21 (2006), Nr. 2, S. 333–352
- [45] CZAJA, S. J.; LEE, C. C.: Information Technology and older adults. In: SEARS, A. (Hrsg.); JACKO, J. A. (Hrsg.): *Human-computer interaction*. Boca Raton and FL: CRC Press, 2009, S. 18–32
- [46] CZERWINSKI, M. ; HORVITZ, E. ; WILHITE, S. : A diary study of task switching and interruptions. In: DYKSTRA-ERICKSON, E. (Hrsg.); TSCHELIGI, M. (Hrsg.): *CHI '04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2004, S. 175–182
- [47] DAVIS, A. C.: The Prevalence of Hearing Impairment and Reported Hearing Disability among Adults in Great Britain. In: *International Journal of Epidemiology* 18 (1989), Nr. 4, S. 911–917. – ISSN 0300–5771
- [48] DEUTSCHES INSTITUT FÜR NORMUNG: *Ergonomie - Körpermaße des Menschen - Teil 2: Werte*. 2005-12. Berlin, 2005
- [49] DEUTSCHES INSTITUT FÜR NORMUNG: *Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten*. Berlin, Januar 1999

- [50] DEUTSCHL, G. ; EGGERT, W. H.; OERTEL, W. H.; POEWE, W. : Parkinson-Krankheit. In: OERTEL, W. H. (Hrsg.); POEWE, W. (Hrsg.); DEUTSCHL, G. (Hrsg.): *Parkinsonsyndrome und andere Bewegungsstörungen*. Stuttgart: Georg Thieme Verlag, 2011. – ISBN 978-3-13-166131-9, S. 35-105
- [51] DEUTSCHL, G. ; RAETHJEN, J. ; KÖSTER, B. : Tremor. In: CEBALLOS-BAUMANN, A. O. (Hrsg.); ABELE, M. (Hrsg.): *Bewegungsstörungen*. Stuttgart: Georg Thieme Verlag, 2005. – ISBN 978-3-13-102392-6, S. 255-286
- [52] DICKINSON, A. ; ARNOTT, J. ; PRIOR, S. : Methods for human-computer interaction research with older people. In: *Behaviour & Information Technology* 26 (2007), Nr. 4, S. 343-352. – ISSN 0144-929X
- [53] DIENER, H.-C. ; WELLER, M. : Tremor. In: GEROK, W. (Hrsg.): *Die innere Medizin*. Stuttgart and New York: Schattauer, 2007. – ISBN 978-3-7945-2222-4, S. 1410-1411
- [54] DIX, A. ; GHAZALI, M. ; GILL, S. ; HARE, J. ; RAMDUNY-ELLIS, D. : Physigrams: modelling devices for natural interaction. In: *Formal Aspects of Computing* 21 (2009), Nr. 6, S. 613-641. – ISSN 0934-5043
- [55] DRECHSLER, R. : *Tübinger Beiträge zur Linguistik*. Bd. 425: *Sprachstörungen nach Schädelhirntrauma: Diskursanalytische Untersuchungen aus textlinguistischer und neuropsychologischer Sicht*. Tübingen: Gunter Narr Verlag, 1997. – ISBN 978-3-8233-5090-3
- [56] DRUZDZEL, M. J.: SMILE: Structural Modeling, Inference, and Learning Engine and GeNie: a development environment for graphical decision-theoretic models. In: *AAAI/IAAI*, 1999, S. 902-903
- [57] EHRHARDT, M. : Parkinson-Syndrom. In: COMBERG, H.-U. (Hrsg.); BARLET, J. (Hrsg.): *Allgemeinmedizin*. Stuttgart and New York: Georg Thieme Verlag, 2004 (Intensivkurs zur Weiterbildung). – ISBN 978-3-13-126814-3, S. 433-435
- [58] EILERS, K. ; NACHREINER, F. ; HÄNECKE, K. : Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. In: *Zeitschrift für Arbeitswissenschaft* 40 (1986), Nr. 4, S. 214-224
- [59] ENGELBRECHT, K.-P. : *Estimating Spoken Dialog System Quality with User Models*. Springer Berlin Heidelberg, 2012 (T-Labs Series in Telecommunication Services). – ISBN 978-3-642-31590-9
- [60] ENGELBRECHT, K.-P. ; KRUPPA, M. ; MÖLLER, S. ; QUADE, M. : MeMo workbench for semi-automated usability testing. In:

Ninth Annual Conference of the International Speech Communication Association, 2008, S. 1662–1665

- [61] ENGELHARDT, M. : Epidemiologie der Arthrose in Westeuropa. In: *Deutsche Zeitschrift für Sportmedizin* 54 (2003), Nr. 6, S. 171–175
- [62] EUROPEAN COMMISSION; EUROSTAT (Hrsg.): *Ageing characterises the demographic perspectives of the European societies*. <http://bookshop.europa.eu/en/ageing-characterises-the-demographic-perspectives-of-the-european-societies-pbKSSF08072/>. Version: 2008 (Statistics in focus)
- [63] EYSEL, U. : Sehen und Augenbewegungen. In: SCHMIDT, R. F. (Hrsg.); LANG, F. (Hrsg.): *Physiologie des Menschen*. Berlin and Heidelberg: Springer Berlin Heidelberg, 2007 (Springer-Lehrbuch), S. 377–420
- [64] FAUST, V. ; ARBEITSGEMEINSCHAFT PSYCHOSOZIALE GESUNDHEIT (Hrsg.): *Depression - Teil 1: Beschwerdebild - Ursachen - Krankheitsverlauf*. <http://www.psychosoziale-gesundheit.net/psychiatrie/depression4a.html>. Version: 03.03.2015 (Psychiatrie heute: Seelische Störungen erkennen, verstehen, verhindern, behandeln)
- [65] FELDMANN, H. ; BRUSIS, T. : *Das Gutachten des Hals-Nasen-Ohren-Arztes*. 7. Stuttgart: Thieme, 2012. – ISBN 978-3-13-160047-9
- [66] FEUERSTACK, S. ; BLUMENDORF, M. ; KERN, M. ; KRUPPA, M. ; QUADE, M. ; RUNGE, M. ; ALBAYRAK, S. : Automated Usability Evaluation during Model-Based Interactive System Development. In: FORBRIG, P. (Hrsg.); PATERNÒ, F. (Hrsg.): *Engineering Interactive Systems 2008* Bd. 5247. Berlin and Heidelberg: Springer Berlin Heidelberg, 2008. – ISBN 978-3-540-85991-8, S. 134–141
- [67] FISK, A. D.: *Designing for older adults: Principles and creative human factors approaches*. 2nd. Boca Raton and FL: CRC Press, 2009. – ISBN 978-1-4200-8055-1
- [68] FRANTZIDIS, C. A.; BAMIDIS, P. D.: Description and Future Trends of ICT Solutions Offered Towards Independent Living: The Case of LLM Project. In: *Proceedings of the 2Nd International Conference on PErvasive Technologies Related to Assistive Environments*. New York and NY and USA: ACM, 2009 (PETRA '09). – ISBN 978-1-60558-409-6, S. 59:1–59:8
- [69] FRIEDMAN-HILL, E. : *Jess in action: Rule-based systems in Java*. Greenwich and CT: Manning, 2003. – ISBN 978-1-930110-89-2

- [70] FU, W.-T. ; PIROLI, P. : SNIF-ACT: A Cognitive Model of User Navigation on the World Wide Web. In: *Human-Computer Interaction* 22 (2007), Nr. 4, S. 355-412
- [71] FUNKE, J. : Intelligenz. In: FUNKE, J. (Hrsg.): *Handbuch der Allgemeinen Psychologie - Kognition* Bd. 5. Göttingen [u.a.]: Hogrefe, 2006. – ISBN 978-3-8409-1846-9, S. 48-56
- [72] GAESE, F. : Intelligenzminderung (F7). In: AMANN, R. (Hrsg.); SCHANZE, C. (Hrsg.): *Psychiatrische Diagnostik und Therapie bei Menschen mit Intelligenzminderung*. Stuttgart [u.a.]: Schattauer, 2007. – ISBN 978-3-7945-2422-8, S. 216-232
- [73] GEARY, D. C.: Evolution menschlicher Kognition. In: FUNKE, J. (Hrsg.): *Handbuch der Allgemeinen Psychologie - Kognition* Bd. 5. Göttingen [u.a.]: Hogrefe, 2006. – ISBN 978-3-8409-1846-9, S. 29-38
- [74] GEDIGA, G. ; HAMBORG, K.-C. : Evaluation in der Software-Ergonomie. In: *Zeitschrift für Psychologie* 210 (2002), Nr. 1, S. 40-57. – ISSN 0044-3409
- [75] GEISSLER, R. : *Die Sozialstruktur Deutschlands: Zur gesellschaftlichen Entwicklung mit einer Bilanz zur Vereinigung*. 7. Wiesbaden: Springer Fachmedien Wiesbaden GmbH, 2013. – ISBN 978-3-531-18629-0
- [76] GENTNER, D. ; STEVENS, A. L.: *Mental models*. Hillsdale and N.J.: Erlbaum, 1983 (Cognitive science). – ISBN 978-0-89859-242-9
- [77] GEORGILA, K. ; WOLTERS, M. ; MOORE, J. D.: Simulating the behaviour of older versus younger users when interacting with spoken dialogue systems. In: *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers*, 2008, S. 49-52
- [78] GIAKOU MIS, D. ; KAKLANIS, N. ; VOTIS, K. ; TZOVARAS, D. : Enabling user interface developers to experience accessibility limitations through visual, hearing, physical and cognitive impairment simulation. In: *Universal Access in the Information Society* 13 (2014), Nr. 2, S. 227-248. – ISSN 1615-5289
- [79] GILB, T. ; FINZI, S. : *Principles of software engineering management*. Wokingham and Eng and Reading and Mass: Addison-Wesley Pub. Co., 1988. – ISBN 978-0-201-19246-9
- [80] GOODMAN, J. : Older Adults' Use of Computers: A Survey. In: *Proc. HCI 2003*, Springer-Verlag, 2003, S. 8-12

- [81] GOODMAN-DEANE, J. ; LANGDON, P. M.; CLARKSON, P. J.; CALDWELL, N. H. M.; SARHAN, A. M.: Equipping designers by simulating the effects of visual and hearing impairments. In: PONTELLI, E. (Hrsg.); TREWIN, S. (Hrsg.): *Assets '07 Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, 2007, S. 241–242
- [82] GREGOR, P. ; NEWELL, A. F.; ZAJICEK, M. : Designing for Dynamic Diversity: Interfaces for Older People. In: *Proceedings of the Fifth International ACM Conference on Assistive Technologies*. New York and NY and USA: ACM, 2002 (Assets '02). – ISBN 1–58113–464–9, S. 151–156
- [83] GREHN, F. : Refraktionsfehler: Brillen, Kontaktlinsen und refraktive Chirurgie. In: GREHN, F. (Hrsg.): *Augenheilkunde*. Berlin and Heidelberg: Springer Berlin Heidelberg, 2012 (Springer-Lehrbuch). – ISBN 978–3–642–11332–1, S. 357–376
- [84] HAERTING, J. ; KLUTTIG, A. ; GREISER, K. ; NUDING, S. ; WERDAN, K. : Kohortenstudie zu Risikofaktoren für Herz-Kreislauf-Krankheiten in einer urbanen älteren ostdeutschen Allgemeinbevölkerung (CARLA-Studie). In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 55 (2012), Nr. 6-7, S. 795–800. – ISSN 1436–9990
- [85] HANNULA, S. : *Hearing among older adults: An epidemiological study*. Oulu, University of Oulu, Diss., 2011
- [86] HANSON, V. L.: Influencing technology adoption by older adults. In: *Interacting with Computers* 22 (2010), Nr. 6, S. 502–509. – ISSN 09535438
- [87] HANSON, V. L.: Technology skill and age: what will be the same 20 years from now? In: *Universal Access in the Information Society* 10 (2011), Nr. 4, S. 443–452. – ISSN 1615–5289
- [88] HASSENZAHL, M. ; MONK, A. : The Inference of Perceived Usability From Beauty. In: *Human-Computer Interaction* 25 (2010), Nr. 3, S. 235–260. – ISSN 0737–0024
- [89] HÄSSLER, F. : *Intelligenzminderung: Eine ärztliche Herausforderung*. Berlin and Heidelberg [u.a.]: Springer, 2011 (Manuale psychischer Störungen bei Kindern und Jugendlichen). – ISBN 978–3–642–12995–7
- [90] HAUG, S. ; BUNDESAMT FÜR MIGRATION UND FLÜCHTLINGE (Hrsg.): *Sprachliche Integration von Migranten in Deutschland*. Nürnberg, Mai 2008 (Integrationsreport)
- [91] HÄUSSLER, B. ; KLEIN, S. ; HAGENMEYER, E.-G. : *Weissbuch Diabetes in Deutschland: Bestandsaufnahme und Zukunftsperspektiven ;*

- 20 Tabellen. 2. Stuttgart and New York and NY: Thieme, 2010. – ISBN 978-3-13-143702-0
- [92] HAWTHORN, D. : Possible implications of aging for interface designers. In: *Interacting with Computers* 12 (2000), Nr. 5, S. 507–528. – ISSN 0953-5438
- [93] HECKMANN, D. ; SCHWARTZ, T. ; BRANDHERM, B. ; SCHMITZ, M. ; WILAMOWITZ-MOELLENDORFF, M. v.: Gumo – The General User Model Ontology. In: HUTCHISON, D. (Hrsg.); KANADE, T. (Hrsg.); KITTLER, J. (Hrsg.); KLEINBERG, J. M. (Hrsg.); MATTERN, F. (Hrsg.); MITCHELL, J. C. (Hrsg.); NAOR, M. (Hrsg.); NIERSTRASZ, O. (Hrsg.); PANDU RANGAN, C. (Hrsg.); STEFFEN, B. (Hrsg.); SUDAN, M. (Hrsg.); TERZOPOULOS, D. (Hrsg.); TYGAR, D. (Hrsg.); VARDI, M. Y. (Hrsg.); WEIKUM, G. (Hrsg.); ARDISSONO, L. (Hrsg.); BRNA, P. (Hrsg.); MITROVIC, A. (Hrsg.): *User Modeling 2005* Bd. 3538. Berlin and Heidelberg: Springer Berlin Heidelberg, 2005. – ISBN 978-3-540-27885-6, S. 428–432
- [94] HEDGES, L. V.: Distribution theory for Glass's estimator of effect size and related estimators. In: *Journal of Educational and Behavioral Statistics* 6 (1981), Nr. 2, S. 107–128
- [95] HEGER, D. ; HOLUBE, i. : Wie viele Menschen sind schwerhörig? In: *Zeitschrift für Audiologie* 49 (2010), Nr. 2, S. 61–70
- [96] HEIDEMANN, C. ; DU, Y. ; SCHUBERT, I. ; RATHMANN, W. ; SCHEIDT-NAVE, C. : Prävalenz und zeitliche Entwicklung des bekannten Diabetes mellitus. In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 56 (2013), Nr. 5-6, S. 668–677. – ISSN 1436-9990
- [97] HEIDEMANN, C. ; DU, Y. ; SCHEIDT-NAVE, C. : *Diabetes mellitus in Deutschland*. Berlin, 2011 (GBE Kompakt)
- [98] HESSE, G. ; LAUBERT, A. : Hörminderung im Alter–Ausprägung und Lokalisation. In: *Deutsches Ärzteblatt* 102 (2005), Nr. 41, S. 2864–2868
- [99] HEUSCHMANN, P. ; BUSSE, O. ; WAGNER, M. ; ENDRES, M. ; VILLRINGER, A. ; RÖTHER, J. ; KOLOMINSKY-RABAS, P. ; BERGER, K. : Schlaganfallhäufigkeit und Versorgung von Schlaganfallpatienten in Deutschland. In: *Aktuelle Neurologie* 37 (2010), Nr. 07, S. 333–340. – ISSN 0302-4350
- [100] HOFFMANN, S. ; VITZTHUM, K. ; MACHE, S. ; SPALLEK, M. ; QUARCOO, D. ; GRONEBERG, D. A.; UIBEL, S. : Multiple Sklerose: Epidemiologie, Pathophysiologie, Diagnostik und Therapie. In: *Praktische Arbeitsmedizin* 17 (2009), Nr. 1, S. 12–18

- [101] HOFMAN, A. ; ROCCA, W. A.; BRAYNE, C. ; BRETELER, M. M. B.; CLARKE, M. ; COOPER, B. ; COPELAND, J. R. M.; DARTIGUES, J. F.; DROUX, A. D. S.; HAGNELL, O. ; HEEREN, T. J.; ENGEDAL, K. ; JONKER, C. ; LINDESAY, J. ; LOBO, A. ; MANN, A. H.; MÖLSÄ, P. K.; MORGAN, K. ; O'CONNOR, D. W.; SULKAVA, R. ; KAY, D. W. K.; AMADUCCI, L. : The Prevalence of Dementia in Europe: A Collaborative Study of 1980–1990 Findings. In: *International Journal of Epidemiology* 20 (1991), Nr. 3, S. 736–748. – ISSN 0300–5771
- [102] HOLM, S. : A simple sequentially rejective multiple test procedure. In: *Scandinavian journal of statistics* (1979), S. 65–70
- [103] HOLZINGER, A. ; SEARLE, G. ; NISCHELWITZER, A. : On Some Aspects of Improving Mobile Applications for the Elderly. In: STEPHANIDIS, C. (Hrsg.): *Universal Access in Human Computer Interaction. Coping with Diversity* Bd. 4554. Berlin and Heidelberg: Springer Berlin Heidelberg, 2007. – ISBN 978–3–540–73278–5, S. 923–932
- [104] HOMMEL, B. ; NATTKEMPER, D. : *Handlungspsychologie. Planung und Kontrolle internationalen Handelns*. Berlin: Springer, 2010 (Springer-Lehrbuch Psychologie). – ISBN 978–3–642–12857–8
- [105] HORN, J. L.; CATTELL, R. B.: Age differences in fluid and crystallized intelligence. In: *Acta Psychologica* 26 (1967), S. 107–129. – ISSN 00016918
- [106] (HRSG.), S. B.: *Bevölkerung Deutschlands bis 2060: 12. koordinierte Bevölkerungsvorausberechnung*. Wiesbaden, 18. November 2009
- [107] (HRSG.), S. B.: *Sozialleistungen: Schwerbehinderte Menschen*. Wiesbaden, 2013 (Fachserie 13 / Reihe 5.1)
- [108] HUBER, W. ; POECK, K. ; SPRINGER, L. : *Klinik und Rehabilitation der Aphasie: Eine Einführung für Therapeuten, Angehörige und Betroffene*. 2. Stuttgart: Thieme, 2013. – ISBN 978–3–13–173862–2
- [109] INSTITUT FÜR DEMOSKOPIE ALLENSBACH (HRSG.); INSTITUT FÜR DEMOSKOPIE ALLENSBACH (Hrsg.): *Allensbacher Marktanalyse Werbeträgeranalyse: Codebuch*. Allensbach, 4. Juli 2013
- [110] JACOB, R. J.: Human-computer interaction: input devices. In: *ACM Computing Surveys (CSUR)* 28 (1996), Nr. 1, S. 177–179
- [111] JAEGGI, S. M.; BUSCHKUEHL, M. ; JONIDES, J. ; PERRIG, W. J.: Improving fluid intelligence with training on working memory. In: *Proceedings of the National Academy of Sciences* 105 (2008), Nr. 19, S. 6829–6833. – ISSN 0027–8424

- [112] JAKOBS, E.-M. ; LEHNEN, K. ; ZIEFLE, M. : *Alter und Technik: Studie zu Technikkonzepten, Techniknutzung und Technikbewertung älterer Menschen*. BoD-Books on Demand, 2008
- [113] JAMESON, A. ; MAHR, A. ; KRUPPA, M. ; RIEGER, A. ; SCHLEICHER, R. : Looking for Unexpected Consequences of Interface Design Decisions: The MeMo Workbench. In: WINCKLER, M. (Hrsg.); JOHNSON, H. (Hrsg.); PALANQUE, P. (Hrsg.): *Task Models and Diagrams for User Interface Design* Bd. 4849. Berlin and Heidelberg: Springer Berlin Heidelberg, 2007. – ISBN 978-3-540-77221-7, S. 279-286
- [114] JANHSEN, K. ; STRUBE, H. ; STARKER, A. ; ROBERT KOCH-INSTITUT (Hrsg.): *Hypertonie*. Berlin, 2008 (Gesundheitsberichterstattung des Bundes)
- [115] JASTRZEMBSKI, T. S.; MYERS, C. ; CHARNESS, N. : A Principled Account of the Older Adult in ACT-R: Age-Specific Model Human Processor Extensions in a Mobile Phone Task. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 54 (2010), Nr. 13, S. 982-986. – ISSN 1071-1813
- [116] JASTRZEMBSKI, T. S.: *The Model Human Processor and the older adults: Validation and error extension to GOMS in a mobile phone task*, The Florida State University, Diss., 2006
- [117] JASTRZEMBSKI, T. S.; CHARNESS, N. : The Model Human Processor and the older adult: Parameter estimation and validation within a mobile phone task. In: *Journal of Experimental Psychology: Applied* 13 (2007), Nr. 4, S. 224-248. – ISSN 1939-2192
- [118] JENSEN, F. V.; NIELSEN, T. D.: *Bayesian networks and decision graphs*. 2. New York: Springer, 2007 (Information science and statistics). – ISBN 978-0-387-68281-5
- [119] JOHN, B. ; JASTRZEMBSKI, T. : Exploration of costs and benefits of predictive human performance modeling for design. In: SALVUCCI, D. D. (Hrsg.); GUNZELMANN, G. (Hrsg.): *Proceedings of the 10th International Conference on Cognitive Modeling, Philadelphia, PA*, 2010, S. 115-120
- [120] JOHN, B. E.; PREVAS, K. ; SALVUCCI, D. D.; KOEDINGER, K. : Predictive human performance modeling made easy. In: *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*, ACM Press, 2004. – ISBN 1-58113-702-8, S. 455-462
- [121] JOHN M. CARROLL; JUDITH REITMAN OLSON: *Mental Models in Human-Computer Interaction: Research Issues About What the User of Software Knows*. National Academies Press, 1987. – ISBN 978-0-309-07801-6

- [122] JUST, M. A.; CARPENTER, P. A.: A capacity theory of comprehension: Individual differences in working memory. In: *Psychological Review* 99 (1992), Nr. 1, S. 122–149. – ISSN 1939–1471
- [123] KAKARLAPUDI, V. ; SAWYER, R. ; STAECCKER, H. : The effect of diabetes on sensorineural hearing loss. In: *Otology & Neurotology* 24 (2003), Nr. 3, S. 382–386
- [124] KANG, N. E.; YOON, W. C.: Age- and experience-related user behavior differences in the use of complicated electronic devices. In: *International Journal of Human-Computer Studies* 66 (2008), Nr. 6, S. 425–437. – ISSN 10715819
- [125] KARAVIDAS, M. ; LIM, N. K.; KATSIKAS, S. L.: The effects of computers on older adult users. In: *Computers in Human Behavior* 21 (2005), Nr. 5, S. 697–711
- [126] KARBE, H. ; KÜST, J. : Rehabilitation von Sprach- und Sprechstörungen. In: NELLES, G. (Hrsg.): *Neurologische Rehabilitation*. Stuttgart: Thieme, 2004 (RRN - Referenz-Reihe Neurologie). – ISBN 978–3–13–136261–2, S. 88–103
- [127] KARRER, K. ; GLASER, C. ; CLEMENS, C. ; BRUDER, C. : Technikaffinität erfassen – Der Fragebogen TA-EG. In: LICHTENSTEIN, A. (Hrsg.): *Der Mensch im Mittelpunkt technischer Systeme*. Düsseldorf: VDI-Verlag, 2009. – ISBN 978–3–18–302922–8, S. 196–201
- [128] KASS, R. : Building a user model implicitly from a cooperative advisory dialog. In: *User Modeling and User-adapted Interaction* 1 (1991), Nr. 3, S. 203–258. – ISSN 0924–1868
- [129] KEMPEN, J. H.; O'COLMAIN, B. J.; LESKE, C. ; HAFFNER, S. M.; KLEIN, R. ; MOSS, S. E.; TAYLOR, H. R.; HAMMAN, R. F.; WEST, S. K.; WANG, J. J.; CONGDON, N. G.; FRIEDMAN, D. S.: The Prevalence of Diabetic Retinopathy Among Adults in the United States: The Eye Diseases Prevalence Research Group. In: *Archives of Ophthalmology* 122 (2004), Nr. 4, S. 552–563. – ISSN 0003–9950
- [130] KENNTNER, G. ; BUHL, B. ; MENZEL, H. : *Karlsruher sportwissenschaftliche Beiträge*. Bd. 1: *Sport, Lebensalter und Gesundheit: Soziologische, leistungsbiographische, anthropometrische und medizinische Untersuchungen an Kraft-, Ausdauer- und Nichtsportlern im Seniorenalter*. Karlsruhe: Universitäts-Verlag Karlsruhe, 2006. – ISBN 3–937300–99–6
- [131] KIERAS, D. E.; MEYER, D. E.: An Overview of the EPIC Architecture for Cognition and Performance With Application to Human-Computer Interaction. In: *Human-Computer Interaction* 12 (1997), Nr. 4, S. 391–438

- [132] KINGSLEY, E. C.; SCHOFIELD, N. A.; CASE, K. : A Computer Aid for Man Machine Modelling. In: *Proceedings of the 8th Annual Conference on Computer Graphics and Interactive Techniques*. New York and NY and USA: ACM, 1981 (SIGGRAPH '81). – ISBN 978-0-89791-045-3, S. 163–169
- [133] KIRCHHOF, B. : Die altersabhängige Makuladegeneration. In: *Deutsches Ärzteblatt* 97 (2000), Nr. 21, S. 1111–1115
- [134] KLAVER, C. C. W.; LEEUWEN, R. van; VINGERLING, J. R.; JONG, P. T. V. M. d.: Epidemiologie. In: HOLZ, F. G. (Hrsg.); PAULEIKHOFF, D. (Hrsg.); SPAIDE, R. F. (Hrsg.); BIRD, A. C. (Hrsg.): *Altersabhängige Makuladegeneration*. Berlin and Heidelberg: Springer, 2004. – ISBN 978-3-662-05681-3, S. 2–18
- [135] KLAVER, C. C. W.; WOLFS, R. C. W.; VINGERLING, J. R.; HOFMAN, A. ; JONG, P. T. V. M. d.: Age-Specific Prevalence and Causes of Blindness and Visual Impairment in an Older Population: The Rotterdam Study. In: *Archives of Ophthalmology* 116 (1998), Nr. 5, S. 653. – ISSN 0003-9950
- [136] KLEIN, R. : The Epidemiology of Diabetic Retinopathy. In: DUH, E. J. (Hrsg.): *Diabetic Retinopathy*. Totowa and NJ: Humana Press, 2008. – ISBN 978-1-934115-83-1, S. 67–107
- [137] KLIEGL, R. ; BALTES, P. B.: Das Janusgesicht des Alters: Über Wachstum und Abbau in Intelligenz und Gedächtnis. In: *Das Gehirn und seine Erkrankungen (I): natürliche und künstliche Intelligenz, Kunstfehler, Regresse, Sterbehilfe* (1987), S. 355–376
- [138] KLÖR, H. U.: Epidemiologie der Adipositas. In: WECHSLER, J. G. (Hrsg.): *Adipositas*. Berlin: Blackwell, 2003. – ISBN 978-3-89412-512-7, S. 65–82
- [139] KOBAYASHI, M. ; HIYAMA, A. ; MIURA, T. ; ASAKAWA, C. ; HIROSE, M. ; IFUKUBE, T. : Elderly user evaluation of mobile touchscreen interactions. In: *Human-Computer Interaction–INTERACT 2011*. Springer, 2011, S. 83–99
- [140] KRÄMER, G. ; BESSER, R. : *Multiple Sklerose: Antworten auf die 111 wichtigsten Fragen ; Expertenwissen verständlich gemacht: Hilfen für einen sicheren Umgang mit der Erkrankung*. 6. Stuttgart: TRIAS, 2006. – ISBN 978-3-8304-3333-0
- [141] KÜFNER, H. ; KRAUS, L. : Epidemiologische und ökonomische Aspekte des Alkoholismus. In: *Deutsches Ärzteblatt* 99 (2002), Nr. 14, S. 936–945
- [142] KUJALA, S. ; KAUPPINEN, M. : Identifying and Selecting Users for User-Centered Design. In: *Proceedings of the Third Nordic*

- Conference on Human-computer Interaction*. New York and NY and USA: ACM, 2004 (NordiCHI '04). – ISBN 1-58113-857-1, S. 297-303
- [143] LAIRD, J. : SOAR: An architecture for general intelligence. In: *Artificial Intelligence* 33 (1987), Nr. 1, S. 1-64. – ISSN 00043702
- [144] LAMPERT, T. ; ROBERT KOCH-INSTITUT (HRSG) (Hrsg.): *Rauchen: Aktuelle Entwicklungen bei Erwachsenen*. Berlin, 2011 (GBE Kompakt)
- [145] LANDIS, J. R.; KOCH, G. G.: The Measurement of Observer Agreement for Categorical Data. In: *Biometrics* 33 (1977), Nr. 1, S. 159. – ISSN 0006341X
- [146] LANG, G. : Diabetische Retinopathie. In: BÖHM, B. O. (Hrsg.); PALITZSCH, K.-D. (Hrsg.); ROSAK, C. (Hrsg.); SPINAS, G. A. (Hrsg.): *Klinische Diabetologie*. Berlin and Heidelberg: Springer Berlin Heidelberg, 2013. – ISBN 978-3-642-59539-4, S. 208-218
- [147] LANG, G. K.: Glaukom. In: ESSER, J. (Hrsg.); LANG, G. K. (Hrsg.): *Augenheilkunde*. Stuttgart: Georg Thieme Verlag, 2014. – ISBN 978-3-13-151555-1, S. 176-205
- [148] LANGE, C. : *Daten und Fakten: Ergebnisse der Studie "Gesundheit in Deutschland aktuell 2009"*. Berlin: Robert-Koch-Institut, 2011 (Beiträge zur Gesundheitsberichterstattung des Bundes). – ISBN 978-3-89606-206-2
- [149] LANGE, C. : *Daten und Fakten: Ergebnisse der Studie "Gesundheit in Deutschland aktuell 2010"*. Berlin: Robert-Koch-Inst, 2012 (Beiträge zur Gesundheitsberichterstattung des Bundes). – ISBN 978-3-89606-213-0
- [150] LEHR, U. : Gerontologie: eine interdisziplinäre Wissenschaft. In: 1987). *Gerontologie: Eine interdisziplinäre Wissenschaft. Gerontologie: Eine interdisziplinäre Wissenschaft. Beiträge zur I. Gerontologischen Woche, Heidelberg* 9 (1987), Nr. 13.6, S. 1986
- [151] LEMMEN, K. : Pathophysiologie der diabetischen Folgeschäden. In: HAAK, T. (Hrsg.); PALITZSCH, K.-D. (Hrsg.): *Diabetologie für die Praxis*. Stuttgart: Thieme, 2012. – ISBN 978-3-13-162191-7, S. 432-439
- [152] LENARZ, T. ; BOENNINGHAUS, H.-G. : *Hals-Nasen-Ohren-Heilkunde*. 14. Berlin and Heidelberg: Springer Berlin Heidelberg, 2012 (SpringerLink). – ISBN 978-3-642-21131-7
- [153] LEFLOW, B. ; PAETOW, K. : Neurologische Erkrankungen. In: EHLERT, U. (Hrsg.): *Verhaltensmedizin*. Berlin and Heidelberg: Springer Berlin Heidelberg, 2013 (Springer-Lehrbuch). – ISBN 978-3-642-55587-9, S. 572-602

- [154] LEWIS, C. ; POLSON, P. G.; WHARTON, C. ; RIEMAN, J. : Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. In: CHEW, J. C. (Hrsg.); WHITESIDE, J. (Hrsg.): *CHI '90 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York and NY and USA: ACM, 1990, S. 235–242
- [155] LINDLEY, S. E.; HARPER, R. ; SELLEN, A. : Designing for Elders: Exploring the Complexity of Relationships in Later Life. In: *Proceedings of the 22Nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction - Volume 1*. Swinton and UK and UK: British Computer Society, 2008 (BCS-HCI '08). – ISBN 978-1-906124-04-5, S. 77–86
- [156] LÜDTKE, U. : *Statische retinale Gefäßanalyse: Evaluation von neuen Indizes und Effekt von Blutdruckanstieg bei jungen Normalprobanden*. München, Ludwig-Maximilians-Universität zu München, Diss., 2009
- [157] MACHAN, C. M.; HRYNCHAK, P. K.; IRVING, E. L.: Modeling the prevalence of age-related cataract: Waterloo Eye Study. In: *Optometry & Vision Science* 89 (2012), Nr. 2, S. 130–136
- [158] MACKENZIE, I. S.: Fitts' Law as a Research and Design Tool in Human-Computer Interaction. In: *Human-Computer Interaction* 7 (1992), Nr. 1, S. 91–139. – ISSN 0737-0024
- [159] MAERCKER, A. : *Alterspsychotherapie und klinische Gerontopsychologie*. 2. Berlin Heidelberg: Springer-Verlag, 2015. – ISBN 978-3-642-54722-5
- [160] MATTLE, H. ; MUMENTHALER, M. : *Neurologie*. 13. Stuttgart and New York and NY: Thieme, 2013. – ISBN 978-3-13-157773-3
- [161] MENSINK, G. ; SCHIENKIEWITZ, A. ; HAFTENBERGER, M. ; LAMPERT, T. ; ZIESE, T. ; SCHEIDT-NAVE, C. : Übergewicht und Adipositas in Deutschland. In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 56 (2013), Nr. 5-6, S. 786–794. – ISSN 1436-9990
- [162] MOFFATT, K. A.: *Addressing age-related pen-based target acquisition difficulties*, The University of British Columbia (Vancouver, Diss., 2010
- [163] MÖLLER, S. : *Quality Engineering: Qualität kommunikationstechnischer System*. Berlin: Springer, 2010. – ISBN 978-3-642-11547-9
- [164] MÖLLER, S. ; ENGLERT, R. ; ENGELBRECHT, K. ; HAFNER, V. ; JAMESON, A. ; OULASVIRTA, A. ; RAAKE, A. ; REITHINGER, N. : MeMo: Towards Automatic Usability Evaluation of Spoken Dialogue Services by User Error Simulations. In: *Proceedings of*

INTERSPEECH 2006, the Ninth International Conference on Spoken Language Processing, 2006, S. 1786–1789

- [165] MÖLLER, S. ; KREBBER, J. ; RAAKE, A. ; SMEELE, P. ; RAJMAN, M. ; MELICHAR, M. ; PALLOTTA, V. ; TSAKOU, G. ; KLADIS, B. ; VOVOS, A. ; HOONHOUT, J. ; SCHUCHARDT, D. ; FAKOTAKIS, N. ; GANCHEV, T. ; POTAMITIS, I. : INSPIRE: Evaluation of a Smart-Home System for Infotainment Management and Device Control. In: *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)* Bd. 5, 2004, S. 1603–1606
- [166] NEUGARTEN, B. L.; NEUGARTEN, D. A.: *The meanings of age: Selected papers of Bernice L. Neugarten*. Chicago: University of Chicago Press, 1996. – ISBN 978-0-226-57384-7
- [167] NEUHAUSER, H. ; SCHIENKIEWITZ, A. ; SCHAFFRATH ROSARIO, A. ; DORTSCHY, R. K. B.-M. ; ROBERT KOCH-INSTITUT (HRSG) (Hrsg.): *Referenzperzentile für anthropometrische Maßzahlen und Blutdruck aus der Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland (KiGGS): 2. erweiterte Auflage*. 2. Berlin, 2013 (Beiträge zur Gesundheitsberichterstattung des Bundes)
- [168] NEWELL, A. ; ARNOTT, J. ; CARMICHAEL, A. ; MORGAN, M. : Methodologies for Involving Older Adults in the Design Process. In: STEPHANIDIS, C. (Hrsg.): *Universal Access in Human Computer Interaction. Coping with Diversity* Bd. 4554. Berlin and Heidelberg: Springer Berlin Heidelberg, 2007. – ISBN 978-3-540-73278-5, S. 982–989
- [169] NEWELL, A. ; SIMON, H. A.: *Human problem solving*. Englewood Cliffs and N.J: Prentice-Hall, 1972. – ISBN 978-0-13-445403-0
- [170] NIEMANN, H. ; GAUGGEL, S. : Störungen der Aufmerksamkeit. In: KARNATH, H.-O. (Hrsg.): *Kognitive Neurologie*. Stuttgart: Georg Thieme Verlag, 2006 (RRN - Referenz-Reihe Neurologie). – ISBN 978-3-13-136521-7, S. 111–125
- [171] NIETHARD, F. U.; BIBERTHALER, P. ; PFEIL, J. : *Orthopädie und Unfallchirurgie*. 7. Stuttgart: Thieme, 2014 (Duale Reihe Orthopädie und Unfallchirurgie). – ISBN 978-3-13-151037-2
- [172] NOLL, H.-H. ; WEICK, S. ; GESIS (Hrsg.): *Schichtenzugehörigkeit nicht nur vom Einkommen bestimmt*. Mannheim, 2011 (Informationsdienst Soziale Indikatoren)
- [173] NORMAN, D. A.: *The design of everyday things*. 1. New York: Basic Books, 2002, c1988. – ISBN 978-0-465-06710-7
- [174] O'COLMAIN, B. J.; MUÑOZ, B. ; TOMANY, S. C.; MCCARTY, C. ; JONG, P. T. V. M. d.; NEMESURE, B. ; MITCHELL, P. ; KEMPEN,

- J. H.; CONGDON, N. G.: Prevalence of Age-Related Macular Degeneration in the United States. In: *Archives of Ophthalmology* 122 (2004), Nr. 4, S. 564–572. – ISSN 0003–9950
- [175] OERTEL, W. H.: Grundlagen der Bewegungserkrankungen. In: OERTEL, W. H. (Hrsg.); POEWE, W. (Hrsg.); DEUTSCHL, G. (Hrsg.): *Parkinsonsyndrome und andere Bewegungsstörungen*. Stuttgart: Georg Thieme Verlag, 2011. – ISBN 978–3–13–166131–9, S. 3–31
- [176] OFFICE, C. S.: *National disability survey 2006*. Dublin: Stationery Office, 2010. – ISBN 978–1–4064–2352–5
- [177] PABST, A. ; KRAUS, L. : Alkoholkonsum, alkoholbezogene Störungen und Trends. Ergebnisse des Epidemiologischen Suchtsurveys 2006. In: *SUCHT - Zeitschrift für Wissenschaft und Praxis / Journal of Addiction Research and Practice* 54 (2008), Nr. 7, S. 36–46. – ISSN 0939–5911
- [178] PERSAD, U. ; LANGDON, P. ; BROWN, D. ; CLARKSON, P. J.: Cognitive Scales and Mental Models for Inclusive Design. In: STEPHANIDIS, C. (Hrsg.): *Universal Access in Human Computer Interaction. Coping with Diversity* Bd. 4554. Berlin and Heidelberg: Springer Berlin Heidelberg, 2007. – ISBN 978–3–540–73278–5, S. 776–785
- [179] POHL, M. ; MEHRHOLZ, J. : Hintergrund Schlaganfall: Häufigkeiten, Risiken, Überleben, Ursachen, Störungen und Chancen. In: MEHRHOLZ, J. (Hrsg.): *Frühphase Schlaganfall*. s.l: Georg Thieme Verlag KG, 2008. – ISBN 978–3–13–151701–2, S. 1–20
- [180] PORTER, J. ; CASE, K. ; MARSHALL, R. ; GYL, D. ; OLIVER, R. Sims n.: 'Beyond Jack and Jill': designing for individuals using HADRIAN. In: *International Journal of Industrial Ergonomics* 33 (2004), Nr. 3, S. 249–264. – ISSN 01698141
- [181] PÜLLEN, R. ; HARLACHER, R. ; PIENKA, L. ; FÜSGEN, I. : Der ältere Patient mit Schlaganfall - Nachbeobachtung nach 18 Monaten. In: *Zeitschrift für Gerontologie und Geriatrie* 32 (1999), Nr. 5, S. 358–363
- [182] QUARANTA, A. ; ASSENNATO, G. ; SALLUSTIO, V. : Epidemiology of hearing problems among adults in Italy. In: *Scandinavian audiology. Supplementum* 42 (1995), Nr. 25, S. 9–13
- [183] RABENBERG, M. ; ROBERT KOCH-INSTITUT (HRSG) (Hrsg.): *Arthrose*. Berlin, 2013 (Gesundheitsberichterstattung des Bundes)
- [184] RADENBACH, J. : *Lexikon Soziale Betreuung: Fachbegriffe der Altenhilfe von A-Z*. Hannover: Schlütersche Verlagsgesellschaft mbH & Co. KG, 2014. – ISBN 978–3–8426–8589–5

- [185] RAETHJEN, J. ; DEUTSCHL, G. : Tremor. In: OERTEL, W. H. (Hrsg.); POEWE, W. (Hrsg.); DEUTSCHL, G. (Hrsg.): *Parkinsonsyndrome und andere Bewegungsstörungen*. Stuttgart: Georg Thieme Verlag, 2011. – ISBN 978-3-13-166131-9, S. 191–228
- [186] REASON, J. T.: *Human error*. Cambridge [England] and New York: Cambridge University Press, 1990. – ISBN 978-0-521-31419-0
- [187] REEVES, M. J.; BUSHNELL, C. D.; HOWARD, G. ; GARGANO, J. W.; DUNCAN, P. W.; LYNCH, G. ; KHATIWODA, A. ; LISABETH, L. : Sex differences in stroke: epidemiology, clinical presentation, medical care, and outcomes. In: *The Lancet Neurology* 7 (2008), Nr. 10, S. 915–926. – ISSN 14744422
- [188] RIEMAN, J. ; YOUNG, R. M.; HOWES, A. : A dual-space model of iteratively deepening exploratory learning. In: *International Journal of Human-Computer Studies* 44 (1996), Nr. 6, S. 743–775. – ISSN 10715819
- [189] RINGELSTEIN, E. B.; NABAVI, D. G.: *Der ischämische Schlaganfall: Eine praxisorientierte Darstellung von Pathophysiologie, Diagnostik und Therapie*. 1. Stuttgart: Kohlhammer, 2007 (Klinische Neurologie). – ISBN 978-3-17-018853-2
- [190] RINGELSTEIN, E. B.; KNECHT, S. : Bluthochdruck und isolierter systolischer Hypertonus – wichtigste Risikofaktoren des Schlaganfalls. In: *Deutsches Ärzteblatt* 103 (2006), Nr. 27, S. 1906–1911
- [191] RITCHIE, K. ; KILDEA, D. : Is senile dementia “age-related” or “ageing-related”? —evidence from meta-analysis of dementia prevalence in the oldest old. In: *The Lancet* 346 (1995), Nr. 8980, S. 931–934. – ISSN 01406736
- [192] ROBERT KOCH-INSTITUT (HRSG); ROBERT KOCH-INSTITUT (Hrsg.): *Gesundheit in Deutschland*. Berlin, 2006 (Gesundheitsberichterstattung des Bundes)
- [193] ROBERT KOCH-INSTITUT (HRSG): *Daten und Fakten: Ergebnisse der Studie “Gesundheit in Deutschland aktuell 2012“: Beiträge zur Gesundheitsberichterstattung des Bundes*. 1. Berlin: RKI, 2014 (Gesundheitsberichterstattung für Deutschland). – ISBN 978-3-89606-222-2
- [194] ROMANO, J. ; KROMREY, J. D.; CORAGGIO, J. ; SKOWRONEK, J. : Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen’s d for evaluating group differences on the NSSE and other surveys. In: *annual meeting of the Florida Association of Institutional Research*, 2006, S. 1–33

- [195] RUMPL, E. : Schädel-Hirn-Trauma. In: BERLIT, P. (Hrsg.): *Klinische Neurologie*. Heidelberg: Springer Medizin, 2006. – ISBN 978-3-540-01982-4, S. 1074-1095
- [196] RUSS, A. : Modeling Visual Attention for Rule-Based Usability Simulations of Elderly Citizen. In: HUTCHISON, D. (Hrsg.); KANADE, T. (Hrsg.); KITTLER, J. (Hrsg.); KLEINBERG, J. M. (Hrsg.); MATTERN, F. (Hrsg.); MITCHELL, J. C. (Hrsg.); NAOR, M. (Hrsg.); NIERSTRASZ, O. (Hrsg.); PANDU RANGAN, C. (Hrsg.); STEFFEN, B. (Hrsg.); SUDAN, M. (Hrsg.); TERZOPOULOS, D. (Hrsg.); TYGAR, D. (Hrsg.); VARDI, M. Y. (Hrsg.); WEIKUM, G. (Hrsg.); HARRIS, D. (Hrsg.): *Proceeding EPCE'11* Bd. 6781. Berlin and Heidelberg: Springer Berlin Heidelberg, 2011. – ISBN 978-3-642-21740-1, S. 72-81
- [197] SAASE, J. L.; ROMUNDE, L. K.; CATS, A. ; VANDENBROUCKE, J. P.; VALKENBURG, H. A.: Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. In: *Annals of the rheumatic diseases* 48 (1989), Nr. 4, S. 271-280
- [198] SACKMANN, R. ; WEYMANN, A. ; HÜTTNER, B. : *Die Technisierung des Alltags: Generationen und technische Innovationen*. Frankfurt and New York: Campus-Verlag, 1994. – ISBN 978-3-593-35177-3
- [199] SALTHOUSE, T. A.: Aging and measures of processing speed. In: *Biological Psychology* 54 (2000), Nr. 1-3, S. 35-54. – ISSN 03010511
- [200] SALTHOUSE, T. A.: What and When of Cognitive Aging. In: *Current Directions in Psychological Science* 13 (2004), Nr. 4, S. 140-144. – ISSN 0963-7214
- [201] SALVUCCI, D. D.; CHAVEZ, A. K.; LEE, F. J.: Modeling effects of age in complex tasks: A case study in driving. In: *In Proceedings of the 26th Annual Conference of the Cognitive Science Society*, 2004, S. 1197-1202
- [202] SARODNICK, F. ; BRAU, H. : *Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung*. 1. Bern: Huber, 2006 (Praxis der Arbeits- und Organisationspsychologie). – ISBN 978-3-456-84200-4
- [203] SASS, A.-C. : Unfallgeschehen in Deutschland: Auswertung des telefonischen Gesundheitssurveys 2004. In: *Deutsches Ärzteblatt* 105 (2008), Nr. 36, S. 604-608
- [204] SAUP, W. : *Alter und Umwelt: Eine Einführung in die ökologische Gerontologie*. Kohlhammer, 1993. – ISBN 978-3-17-012327-4
- [205] SCHAFFER, S. ; SCHLEICHER, R. ; MÖLLER, S. : Modeling Input Modality Choice in Mobile Graphical and Speech Interfaces. In:

- Int. Journal of Human-Computer Studies* 73 (2015), Nr. 3, S. 21–34.
– ISSN 1071–5819
- [206] SCHEIDT-NAVE, C. ; RICHTER, S. ; FUCHS, J. ; KUHLMAY, A. : Herausforderungen an die Gesundheitsforschung für eine alternde Gesellschaft am Beispiel “Multimorbidität“. In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 53 (2010), Nr. 5, S. 441–450. – ISSN 1436–9990
- [207] SCHROETER, K. R.; KÜNEMUND, H. : “Alter“ als Soziale Konstruktion – eine soziologische Einführung. In: ANER, K. (Hrsg.); KARL, U. (Hrsg.): *Handbuch Soziale Arbeit und Alter*. Wiesbaden: VS Verlag für Sozialwissenschaften, 2009. – ISBN 978–3–531–15560–9, S. 393–401
- [208] SCHUHKNECHT, H. F.: Further Observations On the Pathology of Presbycusis. In: *Archives of Otolaryngology - Head and Neck Surgery* 80 (1964), Nr. 4, S. 369–382. – ISSN 0886–4470
- [209] SCHULZ, M. ; SCHMIDT, S. ; ENGELBRECHT, K.-P. ; MÖLLER, S. : Using device models for analyzing user interaction problems. In: *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility ACM*, 2011, S. 303–304
- [210] SCHÜSSLER, G. ; BRUNNAUER, A. : Psychologische Grundlagen psychischer Erkrankungen. In: MÖLLER, H.-J. (Hrsg.); LAUX, G. (Hrsg.); KAPFFHAMMER, H.-P. (Hrsg.): *Psychiatrie und Psychotherapie*. Berlin and Heidelberg: Springer Berlin Heidelberg, 2008. – ISBN 978–3–540–24583–4, S. 227–264
- [211] SCHWARZ, G. : Estimating the Dimension of a Model. In: *The Annals of Statistics* 6 (1978), Nr. 2, S. 461–464. – ISSN 0090–5364
- [212] SEIDEL, M. : Geistige Behinderung und erworbene Hirnschäden - was sind die Unterschiede? In: SEIDEL, M. (Hrsg.): *Menschen mit erworbenen Hirnschäden – (keine) Randgruppe in der Behindertenhilfe ?!* Bd. 28. Berlin: Eigenverlag der DGSG, 2013. – ISBN 978–3–938931–29–5, S. 11–19
- [213] SETTERSTEN; RICHARD, A. ; MAYER, K. U.: The measurement of age, age structuring, and the life course. In: *Annual review of sociology* (1997), S. 233–261
- [214] SHEWANOWN, S. : *The Usability of Digital Library Learning Resources: Developer Perceptions of the Results of an Empirical Automated Usability Evaluation Approach*, University of Georgia, Diss., 2006
- [215] SHNEIDERMAN, B. : *Designing the user interface: Strategies for effective human-computer interaction*. Reading and Mass: Addison-Wesley, 1987. – ISBN 978–0–201–16505–0

- [216] SILVERMAN, B. W.: *Density estimation for statistics and data analysis*. Bd. 26. CRC press, 1986
- [217] SOHN, W. ; JÖRGENSHAUS, W. : Schwerhörigkeit in Deutschland. In: *Zeitschrift für Allgemeine Medizin* 77 (2001), S. 143–147
- [218] SOLGA, H. ; DORNBROSKI, R. ; HANS-BÖCKLER-STIFTUNG (Hrsg.): *Soziale Ungleichheiten in schulischer und außerschulischer Bildung: Stand der Forschung und Forschungsbedarf*. Düsseldorf, 2009 (Arbeitspapier)
- [219] SOUKOREFF, R. W.; MACKENZIE, I. S.: Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. In: *International Journal of Human-Computer Studies* 61 (2004), Nr. 6, S. 751–789. – ISSN 10715819
- [220] STANDKE, W. ; DEUTSCHE GESETZLICHE UNFALLVERSICHERUNG (DGUV) (Hrsg.): *Unfallgeschehen 2011: Statistik*. München, Januar 2013
- [221] STATISTISCHE ÄMTER DES BUNDES UND DER LÄNDER (HRSG.): *Zensus 2011*. <https://ergebnisse.zensus2011.de>. Version: 2011, letzter Abruf: 10.03.2015
- [222] STATISTISCHE BUNDESAMT (HRSG.): *Schüler/innen mit fremdsprachlichem Unterricht*. <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Schulen/Tabellen/AllgemeinBildendeBeruflicheSchulenFremdsprachUnterricht.html>. Version: 2014
- [223] STATISTISCHE BUNDESAMT (HRSG.): *GENESIS-Online Datenbank*. <https://www-genesis.destatis.de/genesis/online/data>. Version: 2015
- [224] STATISTISCHE BUNDESAMT (HRSG.): *Gesundheitsberichterstattung des Bundes: Daten aus Deutschland*. <https://www.gbe-bund.de>. Version: 2015, letzter Abruf: 11.03.2015
- [225] STÖSSEL, C. ; WANDKE, H. ; BLESSING, L. : Gestural Interfaces for Elderly Users: Help or Hindrance? In: KOPP, S. (Hrsg.); WACHSMUTH, I. (Hrsg.): *Gesture in Embodied Communication and Human-Computer Interaction* Bd. 5934. Springer Berlin Heidelberg, 2010. – ISBN 978–3–642–12552–2, S. 269–280
- [226] STREPPPEL, M. ; WALGER, M. ; WEDEL, H. v.; GABER, E. ; ROBERT KOCH-INSTITUT (Hrsg.): *Themenheft 29 Hörstörungen und Tinnitus*. Berlin, Februar 2006
- [227] STUBBE, H. ; WÖLFER, J. : Schädel-Hirn-Trauma beim Erwachsenen. In: *Intensivmedizin up2date* 08 (2012), Nr. 04, S. 253–269

- [228] SUZMAN, R. ; BEARD, J. : *Global Health and Aging (NIH Publication no. 11-7737)*. http://www.who.int/ageing/publications/global_health.pdf. Version: 2011
- [229] TAUSCHE, A.-K. ; MANGER, B. ; MÜLLER-LADNER, U. ; SCHMIDT, B. : Die Gicht als Systemerkrankung. In: *Zeitschrift für Rheumatologie* 71 (2012), Nr. 3, S. 224–230. – ISSN 0340–1855
- [230] TAUSCHE, A.-K. ; JANSEN, T. L. ; SCHRÖDER, H.-E. ; BORNSTEIN, S. R. ; ARINGER, M. ; MÜLLER-LADNER, U. : Gicht - aktuelle Aspekte in Diagnostik und Therapie. In: *Deutsches Ärzteblatt* 106 (2009), Nr. 34-35, S. 549–555
- [231] TAYLOR, H. R. ; KEEFFE, J. E. ; VU, H. T. V. ; WANG, J. J. ; ROCHTCHINA, E. ; PEZZULLO, M. L. ; MITCHELL, P. : Vision loss in Australia. In: *Med J Aust* 182 (2005), Nr. 11, S. 565–568
- [232] THIEME VERLAG (HRSG.): *Endspurt Klinik: Die Inhalte dieses Werkes basieren überwiegend auf dem Kompendium "AllEx - Alles fürs Examen" ; Skript*. Stuttgart: Thieme, 2013. – ISBN 978–3–13–174431–9
- [233] THORNTON, J. ; EDWARDS, R. ; MITCHELL, P. ; HARRISON, R. A. ; BUCHAN, I. ; KELLY, S. P.: Smoking and age-related macular degeneration: a review of association. In: *Eye* 19 (2005), Nr. 9, S. 935–944. – ISSN 0950–222X
- [234] TREWIN, S. ; JOHN, B. ; RICHARDS, J. ; SLOAN, D. ; HANSON, V. ; BELLAMY, R. ; THOMAS, J. ; SWART, C. : Age-specific Predictive Models of Human Performance. In: *CHI '12 Extended Abstracts on Human Factors in Computing Systems*. New York and NY and USA: ACM, 2012 (CHI EA '12). – ISBN 978–1–4503–1016–1, S. 2267–2272
- [235] TUOMILEHTO, J. ; NUOTTIMAKI, T. ; SALMI, K. ; AHO, K. ; KOTILA, M. ; SARTI, C. ; RASTENYTE, D. : Psychosocial and Health Status in Stroke Survivors After 14 Years. In: *Stroke* 26 (1995), Nr. 6, S. 971–975. – ISSN 0039–2499
- [236] UK PROSPECTIVE DIABETES STUDY (UKPDS) GROUP: Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). In: *The Lancet* 352 (1998), Nr. 9131, S. 837–853. – ISSN 01406736
- [237] VARNACCIA, G. ; ROMMEL, A. ; SASS, A.-C. : *Das Unfallgeschehen bei Erwachsenen in Deutschland: Ergebnisse des Unfallmoduls der Befragung "Gesundheit in Deutschland aktuell 2010"*. Berlin: Robert-Koch-Institut, 2013 (Beiträge zur Gesundheitsberichterstattung des Bundes). – ISBN 978–3–89606–217–8

- [238] VASILYEVA, E. ; PUURONEN, S. ; PECHENIZKIY, M. ; RASANEN, P. : Feedback adaptation in web-based learning systems. In: *International Journal of Continuing Engineering Education and Life Long Learning* 17 (2007), Nr. 4, S. 337–357
- [239] VERA, A. ; HOWES, A. ; MCCURDY, M. ; LEWIS, R. L.: A constraint satisfaction approach to predicting skilled interactive cognition. In: DYKSTRA-ERICKSON, E. (Hrsg.); TSCHELIGI, M. (Hrsg.): *CHI '04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2004, S. 121–128
- [240] WADE, D. T.; HEWER, R. L.: Functional abilities after stroke: measurement, natural history and prognosis. In: *Journal of Neurology, Neurosurgery & Psychiatry* 50 (1987), Nr. 2, S. 177–182. – ISSN 0022–3050
- [241] WAHL, H.-W. ; SCHMITT, M. : Psychogerontologische Konzepte des “Alter(n)s”. In: ANER, K. (Hrsg.); KARL, U. (Hrsg.): *Handbuch Soziale Arbeit und Alter*. Wiesbaden: VS Verlag für Sozialwissenschaften, 2009. – ISBN 978–3–531–15560–9, S. 427–438
- [242] WALLING, A. ; DICKSON, G. : Hearing loss in older adults. In: *American family physician* 85 (2012), Nr. 12, S. 1150–1156. – ISSN 0002–838X
- [243] WARNKE, A. : Intelligenzminderung. In: HERPERTZ-DAHLMANN, B. (Hrsg.): *Entwicklungspsychiatrie*. Stuttgart: F. Schattauer, 2008. – ISBN 978–3–7945–2358–0, S. 487–526
- [244] WASCHER, E. ; RINKENAUER, G. ; FALKENSTEIN, M. : A Function Based Approach towards Adaptive Interfaces for Elderly Users. In: HUTCHISON, D. (Hrsg.); KANADE, T. (Hrsg.); KITTLER, J. (Hrsg.); KLEINBERG, J. M. (Hrsg.); MATTERN, F. (Hrsg.); MITCHELL, J. C. (Hrsg.); NAOR, M. (Hrsg.); NIERSTRASZ, O. (Hrsg.); PANDU RANGAN, C. (Hrsg.); STEFFEN, B. (Hrsg.); SUDAN, M. (Hrsg.); TERZOPOULOS, D. (Hrsg.); TYGAR, D. (Hrsg.); VARDI, M. Y. (Hrsg.); WEIKUM, G. (Hrsg.); STEPHANIDIS, C. (Hrsg.): *Lecture Notes in Computer Science*. Berlin and Heidelberg: Springer Berlin Heidelberg, 2009. – ISBN 978–3–642–02706–2, S. 304–311
- [245] WECHSLER, D. : *Wechsler Intelligenztest für Erwachsene: WIE; Übersetzung und Adaption der WAIS-III*. Frankfurt/M: Harcourt Test Services, 2006. – ISBN 978–0–7491–1541–8
- [246] WEICKSEL, J. ; THYLMANN, M. : *Smartphones stärker verbreitet als normale Handys*. 2014
- [247] WEISS, B. ; WECHSUNG, I. ; MARQUARDT, S. : Assessing ICT User Groups. In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*. New York

- and NY and USA: ACM, 2012 (NordiCHI '12). – ISBN 978-1-4503-1482-4, S. 275–283
- [248] WENDT, C. : *Veränderungen verschiedener anthropometrischer Parameter sowie der Körperzusammensetzung im Verlauf des Seniorenalters unter Berücksichtigung der körperlichen Aktivität: Eine Untersuchung im Rahmen der Gießener Senioren Langzeitstudie*. Gießen, Justus-Liebig-Universität, Diss., 2012
- [249] WETTSTEIN, A. ; CONZELMANN, M. : *Checkliste Geriatrie*. 2. Stuttgart: Georg Thieme Verlag KG, 2001. – ISBN 978-3-13-152812-4
- [250] WIEDMANN, J. : *Die Rehabilitation der Schallempfindungsschwerhörigkeit mittels teilimplantierbarem Vibrant Soundbridge System*. Ulm, Universität Ulm, Diss., 2010
- [251] WILD, K. V.; MATTEK, N. C.; MAXWELL, S. A.; DODGE, H. H.; JIMISON, H. B.; KAYE, J. A.: Computer-related self-efficacy and anxiety in older adults with and without mild cognitive impairment. In: *Alzheimer's & Dementia* 8 (2012), Nr. 6, S. 544–552. – ISSN 15525260
- [252] WILSON, G. ; STEWART, C. ; BREWSTER, S. A.: Pressure-based menu selection for mobile devices. In: SÁ, M. d. (Hrsg.); CARRIÇO, L. (Hrsg.); CORREIA, N. (Hrsg.): *MobileHCI '10 Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, 2010, S. 181–190
- [253] WIRTH, A. : *Adipositas-Fibel*. 2. Berlin and Heidelberg: Springer Berlin Heidelberg, 2012. – ISBN 978-3-642-18252-5
- [254] WIRTH, A. ; HAUNER, H. ; RODEN, M. ; PARHOFER, K. ; MAY, M. ; ST. ENGELI; JORDAN, J. ; SCHULZ, R. ; SCHNEIDER, K. T. M.; GROTE, V. ; TEUCHER, B. ; KAAKS, R. : Komorbiditäten. In: WIRTH, A. (Hrsg.); HAUNER, H. (Hrsg.): *Adipositas*. Berlin: Springer, 2013. – ISBN 978-3-642-22855-1, S. 173–258
- [255] WORLD HEALTH ORGANIZATION; DEUTSCHES INSTITUT FÜR MEDIZINISCHE DOKUMENTATION UND INFORMATION (Hrsg.): *ICD-10-WHO Version 2013: Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme 10. Revision Version 2013*. <http://www.dimdi.de/static/de/klassi/icd-10-who/kodesuche/onlinefassungen/htmlamt12013/>. Version: 2013
- [256] YALDIZLI, Ö. ; KAPPOS, L. : Klinische Grundlagen der Multiplen Sklerose. In: STEINLIN EGLI, R. (Hrsg.): *Multiple Sklerose verstehen und behandeln*. Berlin and Heidelberg and New York and NY: Springer, 2011 (Physiotherapie). – ISBN 978-3-642-17633-3, S. 2–9

- [257] ZAHNERT, T. : Differenzialdiagnose der Schwerhörigkeit. In: *Deutsches Ärzteblatt* 108 (2011), Nr. 25, S. 433–444
- [258] ZHANG, X. ; SAADDINE, J. B.; CHOU, C.-F. ; COTCH, M. F.; CHENG, Y. J.; GEISS, L. S.; GREGG, E. W.; ALBRIGHT, A. L.; KLEIN, B. E. K.; KLEIN, R. : Prevalence of Diabetic Retinopathy in the United States, 2005-2008. In: *JAMA* 304 (2010), Nr. 6, S. 649. – ISSN 0098–7484
- [259] ZIEFLE, M. ; WILKOWSKA, W. : Technology acceptability for medical assistance. In: GERHAUSER, H. (Hrsg.); SIEK, K. (Hrsg.); HORNEGGER, J. (Hrsg.); LUETH, T. C. (Hrsg.): *4th International ICST Conference on Pervasive Computing Technologies for Healthcare*, 2010. – ISBN 978–963–9799–89–9, S. 1–9
- [260] ZINK, A. ; MINDEN, K. ; LIST, S. M.: *Gesundheitsberichterstattung des Bundes*. Bd. 49: *Entzündlich-rheumatische Erkrankungen*. Berlin: Robert-Koch-Inst, 2010. – ISBN 978–3–89606–204–8
- [261] ZOMEREN, A. H.; BURG, W. van d.: Residual complaints of patients two years after severe head injury. In: *Journal of Neurology, Neurosurgery & Psychiatry* 48 (1985), Nr. 1, S. 21–28. – ISSN 0022–3050