GESTURAL INTERFACES FOR ELDERLY USERS

- HELP OR HINDRANCE?

vorgelegt von
Christian Stößel (M.Sc.)

von der Fakultät V
Verkehrs- und Maschinensysteme
der Technischen Universität Berlin

et de la Faculté des Sciences, de la
Technologie et de la Communication
de l'Université du Luxembourg

zur Erlangung des akademischen Grades

Doktor der
Naturwissenschaften (Dr. rer.nat.)

Docteur en
Sciences de l'Ingénieur

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr.-Ing. Henning Meyer
Gutachterin:	Prof. Dr.-Ing. Lucienne Blessing
Gutachter:	Prof. Dr. phil. Manfred Thüning
Gutachter:	Prof. Dr. sc. nat. Hartmut Wandke
Gutachter:	Prof. Dr.-Ing. Matthias Rötting
Gutachter:	Prof. Dr. rer. nat. Dieter Ferring

Tag der wissenschaftlichen Aussprache: 3. Juli 2012

PhD-FSTC-2012-19
Berlin 2012
D 83

Abstract

Older adults (60+) are often facing particular difficulties when interacting with technology. The "digital divide" that exists between younger generations, who adopt technology faster, and older generations, who struggle to use electronic devices, is often grounded in technology which is not tailored to the specific needs and abilities of this user group. The normal aging process is accompanied by changes in perceptual, cognitive and motor resources, entailing implications how to design suitable technology for the elderly.

Gesture-based interaction, and in particular surface gesturing on multitouch screens, has become the de-facto standard of controlling mobile devices. While this technology is often attributed a particularly "natural" and "intuitive" way of usage, its aptness for older users in particular has still to be shown. The purpose of this work is to fathom the potential of gesture interaction to facilitate technology interaction for elderly users, and to determine constraints that need to be respected to render this input paradigm senior friendly. To this end, research has progressed in three stages.

Stage one addressed the question whether the demands on motor control and manual dexterity imposed by finger gesturing match the abilities of older adults. In two studies, the effects of gesture complexity, available screen space, and the operation posture of a mobile device on gesture performance were assessed. Though being slower than younger users, older users did not show systematic or substantial problems in accurately reproducing gesture patterns, even on small screens or thumb-only gesture execution. Age-related slowing was found to be reduced compared with other input devices.

Stage two addressed the semantics of gestures, in search of a set of gesture commands which are intuitive to older adults. In two studies, user-derived gestures were filed into a gesture glossary, and subsequently evaluated to derive a list of best matches for 34 typical tasks on mobile devices. Older users favor single-finger gestures, and are more willing to sacrifice efficiency for familiarity.

In stage three, conventional key input was compared with gesture input for older adults, using a functional mobile phone prototype. Older users were not only more efficient and made fewer errors in the gesture mode, they also rated this input mode as more intuitive and better learnable than the key mode, and 72% indicated gesture control as their preferred interaction mode.

Taken together, the empirical results provide first evidence that interacting with electronic devices by means of touch and gesture does not overtax older adults, but on the contrary, might have benefits over conventional key input. Using this technology could reduce existing interaction problems, and ultimately support elderly better in their daily activities through technology better adapted to their needs and abilities.

Kurzfassung

Ältere Menschen (60+) haben oft größere Probleme im Umgang mit elektronischen Geräten. In vielen Fällen ist die Technik nicht an die Bedürfnisse und Fähigkeiten der älteren Nutzer angepasst. Sensorische, kognitive und motorische Fähigkeiten nehmen im Zuge des Alterungsprozesses ab und erschweren die Bedienung von Alltagstechnik, z.B. durch zu kleine Tasten, die meist noch schlecht beschriftet und mit mehreren Funktionen auf einmal belegt sind.

Gestensteuerung, insbesondere die Verwendung von Fingergesten auf Multi-touch Bildschirmen, hat sich über die letzten Jahre zum Quasi-Standard bei der Bedienung von mobilen Geräten entwickelt. Obwohl diese Art der Steuerung als besonders "intuitiv" und "natürlich" vermarktet wird, ist unklar, ob dies insbesondere bzgl. älterer Nutzer wirklich zutrifft. Ziel der vorliegenden Arbeit ist herauszufinden, ob Gestensteuerung dazu beitragen kann, die Interaktion mit Technik für ältere Menschen zu vereinfachen, und unter welchen Randbedingungen dies zutrifft.

In einer ersten Stufe wurde überprüft, ob die präzise Ausführung von Fingergesten ältere Nutzer hinsichtlich ihrer motorischen Fähigkeiten vor Probleme stellt. In zwei Studien wurde hierzu der Einfluss von Gestenkomplexität, Bildschirmgröße, und Gerätehaltung auf die Ausführungsgeschwindigkeit und Genauigkeit getestet. Obwohl Ältere die Gesten langsamer als Jüngere ausführten, ließen sich keine substantiellen Unterschiede in der Ausführungsgenauigkeit feststellen. Geschwindigkeitseinbußen waren zudem geringer als bei anderen Eingabegeräten.

In Stufe zwei wurde dann die Semantik von Fingergesten untersucht mit dem Ziel, Gesten zu identifizieren, die für ältere Nutzer intuitiv nutzbar sind. Hierzu wurden zunächst nutzergenerierte Gesten katalogisiert und in einer zweiten Studie hinsichtlich ihrer Eignung von jüngeren und älteren Nutzern bewertet. Ältere favorisierten generell 1-Finger-Gesten und nahmen häufig eine weniger effiziente Ausführung in Kauf, sofern die Gesten vertraute Zeichen darstellten. Für 34 Aufgaben auf einem mobilen Gerät wurden so potentiell geeignete Gesten für ältere Nutzer identifiziert.

In Stufe drei wurde schließlich eine Gestensteuerung mit konventioneller Tastensteuerung auf einem Handyprototypen verglichen. Es zeigte sich, dass Ältere mit der Gestensteuerung nicht nur schneller waren und weniger Fehler machten, sie bewerteten diese auch als intuitiver und besser lernbar als die Tastenbedienung, und 72% gaben Gesten als ihre bevorzugte Bedienweise an.

Zusammengenommen deuten die Ergebnisse der empirischen Studien darauf hin, dass eine Gerätesteuerung mit Hilfe von Fingergesten älteren Nutzern keine übermäßigen Probleme bereitet sondern sogar Vorteile gegenüber konventioneller Steuerung bieten kann und somit dazu beitragen könnte, bestehende Interaktionsprobleme mit Alltagstechnik zu mindern.

Acknowledgements

This work is the result of my research within the interdisciplinary research group *prometei* (prospective design of human-technology interaction) at the Technische Universität Berlin and in collaboration with the University of Luxembourg. I am indebted to many people inside and outside these groups that supported me during the course of this project work. The research was funded in parts by the DFG (Deutsche Forschungsgemeinschaft) and in parts by the FNR (Fonds National de la Recherche Luxembourg). I wish to thank both organizations for enabling me to carry out this research.

First and foremost, I would like to express my most special thanks to Prof. Dr.-Ing. Lucienne Blessing, who encouraged me to pursue this line of research, and supported me throughout the ups and downs of this PhD project. Despite being geographically separated, she dedicated lots of her time to ensure regular feedback meetings, mutual visits, and last but not least struggled through the first version of this thesis. The feedback from her was always very insightful and this work profited a lot from it. Invaluable feedback was also regularly given by Prof. Dr. Hartmut Wandke and Prof. Dr. Manfred Thüring. They both were always there for me when I needed advice, and brought up many interesting discussion points, both in personal feedback meetings and their respective colloquia. Furthermore I am also indebted to Prof. Dr. Dieter Ferring and Prof. Dr.-Ing. Matthias Rötting, who showed great interest in my field of research and didn't hesitate to join the board of scientific supervisors. I would also like to thank Prof. Dr. Henning Meyer for chairing the defense.

Most beneficial for this research were the working conditions and positive atmosphere I found within the research group *prometei*. My colleagues helped me innumerable times, both through scientific exchanges and moral support. Since I consider them all as friends rather than colleagues, it is difficult to single out a few, but special mentioning deserve Jens Nachtwei, Saskia Kain, Sebastian Welke, Stefanie Huber, Susanne Dannehl, Simone Schmid, Nina Gerard, Julian Adenauer, Rebecca Wiczorek, Torsten Günzler and Janna Protzak. I am particularly indebted to my "sister" in research cluster 11, Anna Pohlmeier, who stood by me throughout all joyful and discouraging moments of this project, and always had wise and encouraging advice for me. The great working conditions in *prometei* were not least enabled by Sandra Widera, Karin Scherrinsky-Pingel, Steffen Hartwig and Nikolaus Rötting, who were always supportive in administrative and computational issues.

In addition I would like to thank Michael Schwarz for his help with programming for Study 1B, and Andrea Schnock for her assistance in programming the functional prototype for study 3. Christine Sturm deserves my gratitude for her support in the preparation and analysis of studies as well as the acquisition of participants. I highly appreciated the inspiring collaboration with Jörn Hurtienne on the intersection of

gesture interfaces and image schemas, leading to common supervision of students and publications.

Furthermore, I would like to acknowledge Sylvia Le Hong and Dan Mauney for the inspiration to study 2A, and their help in designing the mockup. My thanks go also to Jurek Kirakowski, who allowed me to use the SUMI questionnaire for the usability assessment in study 3. I owe gratitude also to Stefanie Huber, Simone Schmidt, Lucienne Blessing and Manfred Thüring, for reading parts or entire previous versions of this thesis, providing precious feedback on how to improve it.

Last but not least, I want to express my deep gratitude to my parents, who were always a great source of support throughout the whole period of my PhD project, and especially during the most arduous task of writing down the thesis. I am particularly thankful also for my mother's effort in proof-reading the manuscript.

Table of Contents

1	Introduction.....	1
1.1	Aim and Focus of the Thesis.....	2
1.2	Chapter Overview.....	3
2	Older Users and Technology.....	6
2.1	Who is the „Older User“?.....	7
2.1.1	Age Bands.....	7
2.1.2	Technology Generations.....	8
2.1.3	Typologies.....	10
2.1.4	Terminology Used.....	10
2.2	Physiological and Psychological Changes with Age.....	11
2.2.1	Sensation and Perception.....	12
2.2.2	Cognition.....	13
2.2.3	Motor Control.....	16
2.3	Attitudes and Anxiety Towards Technology.....	17
2.4	Technology Acceptance and Adoption.....	18
2.5	Problems during Technology Interaction.....	22
2.5.1	Mobile Interaction of Older Adults.....	22
2.5.2	Benefits and Drawbacks of Different Input Types.....	25
2.6	Designing for Older Adults.....	27
2.7	Chapter Summary.....	29
3	Gesture Control.....	31
3.1	What Is a „Gesture“?.....	32
3.1.1	Gestures in Human-Human Communication.....	32
3.1.2	Gestures in Human-Computer Communication.....	34
3.1.3	Own Definition of Gesture.....	35
3.2	Classification of Gestures for HCI.....	36
3.2.1	Existing Classification Approaches.....	38
3.2.2	A New Classification Approach.....	39
3.3	Technological Aspects of Gesture Control.....	41
3.3.1	Framing Touch Technology.....	42
3.4	Chapter Summary.....	46
4	Gestural Interaction for Elderly Users.....	47

4.1	Existing Solutions	47
4.2	Potential Benefits and Drawbacks of Gesture Interaction.....	50
4.2.1	Potential Benefits.....	51
4.2.2	Potential Drawbacks.....	53
4.3	Research Agenda	55
4.3.1	Research Questions.....	56
4.3.2	Structure of the Research	58
5	Gesture Syntax and Motor Demands	60
5.1	Motor Control and Aging	60
5.2	Display Size and Device Operation Posture	62
5.2.1	Size of the Gesture Space	63
5.2.2	Operation Posture of the Device.....	63
5.3	Hypotheses.....	64
5.4	Methods Study 1A.....	65
5.4.1	The Gesture Set.....	66
5.4.2	Participants and Apparatus.....	66
5.4.3	Procedure	67
5.4.4	Design.....	68
5.4.5	Data Analysis.....	69
5.5	Results Study 1A	70
5.5.1	Sample Description.....	70
5.5.2	Gesture Performances Measures	71
5.5.3	Subjective Gesture Evaluation	75
5.6	Discussion Study 1A	77
5.6.1	The Influence of Space and Complexity on Gesture Performance	77
5.6.2	The Influence of Age Group on Gesture Performance.....	80
5.7	Methods Study 1B	81
5.7.1	The Gesture Set.....	81
5.7.2	Participants and Apparatus.....	82
5.7.3	Procedure	83
5.7.4	Design.....	85
5.7.5	Data Analysis.....	85
5.8	Results Study 1B	86
5.8.1	Sample Description.....	87
5.8.2	One-finger Gesture Performance.....	87
5.8.3	Two-finger Gesture Performance	90
5.8.4	Subjective Gesture Evaluation	92
5.9	Discussion Study 1B.....	95
5.9.1	The Influence of Posture and Difficulty on Gesture Performance	95

5.9.2	The Influence of Age Group on Gesture Performance	97
5.10	Conclusions: Age effects on Gesture Execution Performance	98
6	Gesture Semantics: Which Gestures are Intuitive to Older Users?	101
6.1	What is meant by „Intuitive“?	101
6.1.1	Definition of Intuitive Use	101
6.1.2	Intuition vs. Prior Experience	103
6.2	Participatory Design: Involving the User	104
6.3	Study 2A: User-defined Gestures for Mobile Interaction.....	105
6.4	Methods Study 2A.....	106
6.4.1	Participants	106
6.4.2	Apparatus	106
6.4.3	Procedure	108
6.4.4	Tasks	108
6.4.5	A Gesture Description Language	110
6.4.6	Gesture Analysis	111
6.5	Results Study 2A	113
6.5.1	Types of Gestures.....	113
6.5.2	Number of Fingers and Gesture Complexity	114
6.5.3	Agreement Score	115
6.5.4	Matching Gestures and Tasks	116
6.5.5	Influence of Prior Experience vs. Influence of Age.....	121
6.6	Discussion Study 2A	123
6.6.1	Influence of the Paper Mockup.....	123
6.6.2	Influence of the Analysis Method	125
6.6.3	Scope of the Results	125
6.7	Study 2B: Evaluating a User-defined Gesture Set	126
6.8	Methods Study 2B	127
6.8.1	Tasks and Gestures	127
6.8.2	Participants	128
6.8.3	Apparatus and Procedure	128
6.9	Results Study 2B	130
6.9.1	Gesture Evaluation across Age Groups.....	130
6.9.2	Agreement.....	131
6.9.3	Persisting and Differing Effects across the Studies.....	134
6.9.4	General Criteria for Suitable Gestures	136
6.10	Discussion Study 2B.....	137
6.10.1	Consistency with Study 2A	138
6.10.2	Divergence from Study 2A	139
6.10.3	Age Group Bias	140

6.10.4 A Revised Gesture Set	142
6.11 Conclusion: Intuitive Gestures for Older Users.....	143
7 Gesture Pragmatics: A Comparison of Gesture and Key Control	147
7.1 Development and Description of the Functional Prototype	149
7.1.1 System.....	149
7.1.2 Screen Design and Functionalities	150
7.1.3 Interaction	152
7.2 Methods Study 3.....	153
7.2.1 The Gesture Set.....	153
7.2.2 Apparatus	154
7.2.3 Design.....	155
7.2.4 Procedure	156
7.2.5 Participants	157
7.2.6 Previous Experience	158
7.3 Results Study 3	160
7.3.1 Comparison of Input Mode.....	160
7.3.2 Task Effects	163
7.3.3 Gesture Usage Patterns	167
7.3.4 Influence of Previous Knowledge and Experience	174
7.4 Discussion Study 3	176
7.4.1 Advantages of Gesture Control over Key Control.....	176
7.4.2 Learning effect.....	177
7.4.3 Limitations of the Approach	178
7.4.4 Usage Patterns	183
7.4.5 Influence of Prior Experience	185
7.5 Conclusions : Touch & Gesture outperforms Key Interaction.....	185
8 General Discussion and Outlook	187
8.1 Summary of Results	187
8.1.1 Stage 1: Age Effects on Gesture Execution Performance:	187
8.1.2 Stage 2: Intuitive Gestures for Older Users.....	189
8.1.3 Stage 3: Evaluation: Gesture Control Outperforms Key Control	190
8.2 Limitations of the Research Approach.....	192
8.2.1 Sample Structure	192
8.2.2 Confounding factors.....	194
8.2.3 Methodological Limitations	196
8.2.4 Limitations of the Gesture Set and Prototype.....	198
8.3 Main Contributions.....	200
8.3.1 Theoretical Contribution.....	201
8.3.2 Methodological Contribution.....	202

8.3.3 Empirical Contribution	203
8.4 Practical Implications for Inclusive Gesture Control Design.....	204
8.5 Future Directions.....	205
References	208
List of Abbreviations	233
Appendix	234
A.1 Overview of Mobile Interaction Problems	234
A.2 Study 1A: Gesture Performance on a Stationary Screen	236
A.2.1 Questionnaire Data	236
A.2.2 Gesture Performance Analysis.....	236
A.3 Study 1B: Gesture Performance on a Mobile Device	239
A.3.1 Sample Characteristics	239
A.3.2 Gesture Performance Analysis: One Finger Gestures	239
A.3.3 Gesture Performance Analysis: Two Finger Gestures.....	243
A.4 Study 2A: A User-derived Gesture Set.....	245
A.4.1 Elements of the Gesture Description Language	245
A.4.2 Analysis of Gesture Types	246
A.4.3 Correlation Analysis.....	247
A.5 Study 2B: Evaluations of the Gesture Set.....	250
A.5.1 Analysis of Type, Number of Fingers, and Complexity	250
A.5.2 Analysis of Individual Gestures	251
A.6 Study 3: Gesture Interaction on a Functional Prototype	255
A.6.1 Interaction Steps.....	255
A.6.2 Analysis of Input Mode and Order of Interaction	255
A.6.3 Analysis of Input Mode and Task	256
A.6.4 Analysis of Questionnaire Data	257
A.6.5 Analysis of Prior Experience	258
A.7 Excursus: A Brief History of Multi-Touch Technology	259
A.8 Study Materials.....	261

1 Introduction

In the past few decades we have witnessed unprecedented developments in computer and information technologies. In today's society, technology has become an integral part of work, education, healthcare, communication and entertainment (Czaja & Lee, 2007). Especially the advent of personal computing, the internet and abundance of mobile services dramatically changed the way we communicate, we inform ourselves, we shop, we navigate, we access media and organize our daily life. All these activities are mediated by technology, providing great opportunities to those who can properly use it. However, despite rising rates of technology adoption among older adults in recent years (Charness & Boot, 2009; Zickuhr, 2010), many elderly still struggle to use technologies that have become an integral part of our daily living, such as computers, mobile phones, TV and Hi-Fi systems, ticket machines etc. (Fisk, Rogers, Charness, Czaja, & Sharit, 2004). A wealth of studies has documented the particular difficulties that older users are facing when interacting with technology across a wide range of devices and tasks (e.g. Charness, Bosman, & Elliot, 1995; Jastrzembski, Charness, Holley, & Feddon, 2005; Kelley & Charness, 1995; Kim, et al., 2007; Kurniawan, 2008; McLaughlin, Rogers, & Fisk, 2003; Pak, McLaughlin, Lin, Rogers, & Fisk, 2002; Tang & Peng, 2006; Ziefle & Bay, 2005).

Nevertheless, it is a misperception that older adults are less open or more averse towards technology (O'Hanlon, Bond, Knapp, & Carragher, 2010; Olson, O'Brien, Rogers, & Charness, 2011). Overall, the available data indicates that older people are not "technophobic" and actually willing to use technology (Czaja & Lee, 2007). Rather, they are facing technology which is not adapted to their specific needs and abilities. Many of the reported problems can be directly linked to the relatively well-understood changes in perception, cognition and psychomotor performance that accompany the normal aging process (Fisk, et al., 2004; Mitzner, Mayhorn, Fisk, & Rogers, 2006). Yet, up to now, the majority of applications and devices are designed without considering the demands of this users group. Older users are bypassed regarding the access to technology, and a digital divide still exists between younger and older adults (Czaja & Hiltz, 2005). This is particularly detrimental, since it puts them at a disadvantage in terms of their ability to live independently, to profit from modern healthcare systems, to retrieve important information, or to communicate with peers and thus help to overcome isolation. The exigency of this situation becomes particularly evident if one considers the dramatic increase of older adults in western societies over the upcoming decades. For example in the U.S., the older

population (age 65 and above) is projected to be twice as large in 2030 as compared to the year 2000 (Federal Interagency Forum on Aging Related Statistics, 2010).

What can be done then to alleviate technology use by older adults? Considering the fact that a great deal of the observed problems arise already at the interface level (Fisk, et al., 2004, e.g. screens and fonts which are too small, buttons which are spaced too narrowly, a lack of sufficient feedback etc.), the design of appropriate interfaces lies in the center of attention of this work.

The way how people interact with technical devices has changed considerably over the last years in unforeseen ways, since devices like the Apple iPhone or Nintendo's Wii gaming console have lifted gesture interaction from niche products and prototypes to the mass consumer market. Other manufacturers quickly adopted this novel interface paradigm of interacting directly with fingers or hands through meaningful gestures. As a result, most modern handheld devices offer a touch & gesture interface, and the technology has spread also to tablets, laptop and desktop computers as well as a broad range of other devices (Saffer, 2008). Even though gestural interfaces offer some obvious advantages, such as the possibility of direct manipulation (Shneiderman, 1983), and have been indefatigably marketed as being very "intuitive" and "natural" to use, they also have been prominently criticized as "a step backward in usability" (Norman & Nielsen, 2010, p. 1). Given the increasing dissemination of touch and gesture control to everyday technology and their promise of easy and intuitive use, they lend themselves as prime candidates to facilitate technology interaction also, and particularly, for older adults. However, research is still lacking which systematically investigates how older adults interact with a gesture-controlled device, to which extent the demands of this input type match the abilities of the older adults, and whether it has indeed the potential to reduce the interaction problems observed with conventional interaction paradigms.

1.1 Aim and Focus of the Thesis

The basic aim of this thesis is to find an answer to the question whether gestural interaction provides an advantage for older users, or whether the drawbacks outweigh potential benefits. With respect to the older adults, the focus was put on the *younger-old* (Fisk, et al., 2004), corresponding to an age band of approximately 60 to 75. Regarding gestural interfaces, the thesis concentrates on two-dimensional finger gestures on touch-sensitive surfaces, and their application on primarily mobile devices. In order to fathom the potential of gesture interaction for older adults, a four stage research agenda was pursued:

Two studies were conducted in the first stage with the goal of investigating whether older adults meet the requirements of gesture technology regarding motor control and manual dexterity. In other words, it needed to be clarified in the first place whether problems arise during the manual execution of certain gesture patterns, and how the gesture syntax could possibly be adapted to minimize them. In stage two, the focus was shifted toward the semantics of gestures. The question was addressed which gesture patterns might be especially intuitive to use for older adults, and whether those differ from gestures designed for a younger and tech-savvy user group. Two studies were carried out to this end, resulting in a gesture set of potentially suitable gestures for older users, which were subsequently evaluated. Stage three comprises the conception and implementation of a functional prototype, while in stage four gesture interaction was evaluated with the mobile phone prototype. A final study was conducted which compared the gesture interaction paradigm with conventional key interaction. The benefits and drawbacks were quantified with usability metrics, but also aspects of learning, memorability and acceptance were assessed.

Taken together, the present research takes a holistic stance in addressing some physical, cognitive and behavioral aspects of the older users, and aligns them with syntactic, semantic and pragmatic properties of the gesture interface. On the basis of the combined insights from the empirical investigations and theoretical considerations, recommendations are derived for a system designer, providing some guidance as to whether gesture input should be considered for inclusively designed devices, and which boundary conditions should be respected in order to ensure ease of use for older adults.

1.2 Chapter Overview

Chapter 2 defines what is meant by older users and how they can be classified. The most important age-related changes in sensory, cognitive and psychomotor resources are outlined, and related research concerning older adults' attitudes and anxiety towards technology is summarized. Several models of technology acceptance and adoption are presented and discussed in the context of older users and gesture technology. The chapter continues by accumulating evidence for actual problems that older users are facing with technology, and in particular with mobile phones. Furthermore, different input paradigms are compared with respect to their appropriateness for older users on the basis of existing literature. The chapter concludes by pointing out similarities and differences between design philosophies which already strive to accommodate the particular needs of an older user group.

Chapter 3 describes the theoretical and technical aspects of gesture interaction. It sets out by defining what is meant by “gesture”, and discusses several definitions and classification approaches in both human-human and human-machine interaction. A working definition for gestures in Human-Computer Interaction (HCI) and a novel classification system are presented. The technical sides of a gestural interface are subsequently addressed by framing key terms related to touch and gesture technology. The vast space of different gesture interaction technologies is narrowed down to a specific class of handheld devices operated by finger gestures, which will be subject to subsequent empirical evaluation.

Chapter 4 brings together the characteristics of the older adults and the characteristics of gesture-based interaction. First, a literature overview on the sparse examples of existing touch & gesture devices for older adults is given, before the potential strengths and weaknesses of the technology with respect to this specific user group are juxtaposed. Based on these theoretical assumptions, the research questions are derived and the research agenda is laid out in greater detail. Dimensions of gesture semiotics are aligned with layers of human behavior, resulting in four sequential stages of the research plan. Parts of this chapter (Section 4.2) have been published in a similar way previously in Stöbel and Blessing (2009).

Chapter 5 presents two studies (1A, 1B) investigating gesture interface parameters such as the complexity of pattern, the size of the gesturing space, the number of fingers used for drawing, or the operation posture of a mobile device, and their effects on gesture execution performance. The comparison with a younger control group highlights possible problems specific to the older users. Gesturing performance is measured primarily with respect to errors, speed and accuracy of the pattern retracing. Hypotheses are derived from the literature on motor control and aging, as well as input device studies with older adults. First, the methods, results, and discussion are given for study one, before study two is described in an analogous manner. The chapter concludes with a general discussion, pointing out differences and similarities between both studies. Parts of this chapter have been published in a highly similar way previously in Stöbel, Wandke & Blessing (2009, 2010).

Chapter 6 first provides a theoretical account of the concept of *intuitive* use, before the *participatory design* approach is explained and justified. Subsequently, two studies are presented which were designed to find out which gestures are intuitive to older adults, and whether they differ from younger adults in this respect. In order to classify the user-derived gestures in study 2A, a structured gesture description language is established, and a gesture glossary is presented, containing a large number of possible gestures for common tasks on a mobile device. Group differences are analyzed regarding type of gestures, complexity of gestures, fingers used and participant agreement. Study 2B re-evaluates the potentially best solutions from the previ-

ous study, and results in a list of best matching gestures for a certain task and age group. Consistent and divergent results with the previous study are discussed, as is the influence of prior experience with technology. The theoretical considerations on intuitive use and prior experience can also be found in slightly abridged form in Stöbel & Blessing (2009). Parts of this chapter have been published previously in similar ways in Stöbel & Blessing (2010b, Sections 6.4-6.6), as well as Stöbel & Blessing (2010a, Sections 6.8-6.9).

Chapter 7 comprises the implementation and evaluation stages of the project. First, the development and the interaction characteristics of the functional mobile phone prototype are described. After that a study is presented in which gesture control of a mobile device is compared to conventional key control with regard to their usability for older adults. Results of the behavioral and subjective assessment are reported and discussed. Furthermore, the influence of task, gesture type and previous technology experience are discussed.

Chapter 8 summarizes and integrates the results from the three empirical stages of the project. Subsequently, general limitations of the research approach are discussed. In particular, the boundaries of generalization regarding the sample, the cross-sectional and quasi-experimental approach, as well as the gesture set and prototype are pointed out. The chapter continues by elaborating the main contribution of the work, with respect to theoretical, methodological and empirical advances in the field. The major findings are compiled into practical design guidelines. Lastly, an outlook is provided where open questions and future directions are laid out.

2 Older Users and Technology

Our societies are getting older. The advances in medicine and living standard in the 20th century have resulted in a substantial increase in the number of older people in Germany, the U.S., and most other developed countries in the world. Over the next decades, the number of older people will further increase dramatically, since the so-called *baby boomers* (those born between 1946 and 1964) are starting to turn 65 in 2011. In the U.S. the older population (age 65 and above) is projected to be twice as large in 2030 as compared to the year 2000, growing from 35 million to 72 million and representing nearly 20% of the total U.S. population (Federal Interagency Forum on Aging Related Statistics, 2010). In Germany, this development is even more pronounced: within the next 30 years, more than one third of the total population will be aged 65 and above (Statistisches Bundesamt Deutschland, 2006). In comparison with 48 European countries, Germany is taking the lead in terms of the projected proportion of the population aged 65+. The same is true for the old age dependency ratio, which gives the quotient of people over 65 and people between 15 and 64, reaching about 50% in 2030 (Vienna Institute of Demography, 2010). In fact, as of 2010, Germany has the world's second oldest population (after Japan), and its old-age dependency ratio is second only to Japan and Italy (United Nations, 2010).

These demographic changes have profound structural, societal and economic implications, and challenge policymakers, families, businesses and healthcare providers alike. Older generations need to be included as active and integral pillars of our society instead of being marginalized. They should remain active members of the work force as long as possible, since the traditional assumption that retirement equals the worker's final exit from the labor force does not hold true any longer. The required transition of society can only be successful if huge efforts are made on various levels to foster independence of this age group, from more flexible employment arrangements, remote services in care giving (*telecare*), support of independent living (*ambient assisted living* - AAL), access to information, access to transportation (*accessibility*), to specific communication services and devices as well as educational offers (*life-long learning*). Technology is believed to play a key role in all these fields. However, only if technology is adapted to the older adults' knowledge, needs, and abilities, can it successfully contribute to their individual well being, and help to meet the challenges of an aging society in general.

This chapter sets out framing the term *older user*, giving an overview of different taxonomies and typologies of older adults, before specifying how the term is used in the remainder of this thesis (Section 2.1). A great deal of research has already investi-

gated the numerous cognitive and physiological changes that accompany the normal aging process, of which a brief overview is presented here (Section 2.2). Subsequently, also the attitudes and anxieties that many older people show towards technology are reviewed (Section 2.3), leading to models of technology acceptance and technology adoption among older adults (Section 2.4). These general characteristics of an aging user are followed by specific age-related problems that occur during the interaction with technology (Section 2.5). On the basis of these observations, a variety of different input techniques is discussed, pointing out their relative benefits and shortcomings with respect to older users. Finally, different design philosophies are briefly portrayed which strive to develop products that respect the needs and abilities of older adults (Section 2.6).

2.1 Who is the „Older User“?

While there is no denying the fact that societies as a whole are getting older, it is much more difficult to specify on an individual level, what we actually mean by *being old*: when do we start classifying somebody as *old*, and which characteristics do we associate with that person? Is there something like a *typical* older user, or would it make much more sense to further break up the older population into more coherent subgroups? Various classification approaches exist, which subdivide the older population according to *age*, *generation* or *type*.

2.1.1 Age Bands

When does *old age* start? There are no agreed definitions of this term. Its meaning and connotation have considerably changed in the course of history and still vary greatly across cultures and societies. Kirk (1992) stresses the point that the defining criteria of the *elderly population*, such as temporal age limits, are constantly changing and follow political and economic needs. According to the World Health Organization, “most developed world countries have accepted the chronological age of 65 years as a definition of ‘elderly’ or older person”, but at the same time the authors acknowledge that this concept is inappropriate for Africa (WHO, 2010, p. 1). The United Nations make the cut for people belonging to the older population at the age of 60, and introduce further subgroups of the “oldest old” (persons aged 80 years and over) and the “centenarians” (persons aged 100 years and over), but provide no motivation for these thresholds (United Nations, 2009). The European Commission also proposes a division of the elderly population into age bands, but this time the differentiation

encompasses *older workers* (55-64), *older people* (65-79) and *frail old people* (80+) (Jané-Llopis & Gabilondo, 2008).

Clearly, the definition of *old age* and the onset of respective age bands depends on the underlying dimensions, of which Glascock and Feinman (1980, cited from WHO, 2010) have identified three main categories: *chronology*, *change in social role* and *change in capabilities*. While in the 19th century *old age* was still mostly defined in terms of physiological changes, nowadays the predominant distinction is the change in social role, e.g. the age of retirement. However, Czaja and Moen (2004) have noted that the traditional assumption of retirement being equivalent to definitely cutting the cord from the labor force at a certain age is increasingly becoming obsolete, since older people increasingly engage in flexible work models, oscillating between paid or unpaid work and retirement.

Looking at the definitions of *older adults* in the human factors and HCI literature, we encounter again different suggestions. Charness (2008, p. 549) claims that “a distinction that has proven useful in the general aging literature is to consider chronological age bands such as the young-old (65-74), middle-old (75-84), and old-old (85+)”. From a more work-centered view, Czaja and Moen (2004) see value in distinguishing the 50s to early 70s, as compared to the late 70s to 90s, and regret at the same time that most studies on aging and technology only group people as 65+ or even 55+. Fisk, Rogers, Charness, Czaja & Sharit (2004) suggest a division into what they term the “younger-old”, ranging in age from approximately 60 to 75, and the “older-old”, comprising those individuals beyond 75 years. At the same time, the authors concede that researchers and practitioners rely on chronological age as an index to define older adults only in lack of a more complete understanding of age-related changes, and that aging in reality is a multi-faceted concept, which can be broken up on at least three different dimensions: *biological*, *psychological* and *social*.

Taken together, there is no widely agreed upon age which qualifies an individual to be referred to as “old”, and the meaning of “old” changes across time and cultures. Most classifications use an age of 55, 60, or 65 as lower threshold for the older population. Consensus exists upon the value of further splitting up the group of older adults into age bands. However, the number of sensible groupings and their respective age boundaries is still unclear and disputed and very much depends on the purpose of grouping, i.e. the dimensions that are considered.

2.1.2 Technology Generations

A completely different approach to classifying older users is thinking in terms of *technology generations*. According to Docampo Rama (2001), “a *generation* has been defined as a range of birth cohorts that display similar behavior, and share norms or

values that are based on common sociological environments during their formative years”. The formative period is generally regarded to span the years between 10 and 25, which are believed to be most critical for acquiring norms, values and skills that influence future behavior (Docampo Rama, 2001). For example, using this concept, Becker (2000) clusters cohorts using discontinuous events (e.g. wars or economic breakover) as separations into five generations: the pre-war generation (1910-1920), the silent generation (1930-1945), the baby-boom generation (1946-1954), the lost generation (1955-1969), and generation X (born 1970 or later). In a similar way, Sackmann and Weymann (1994) tried to separate generations that differ along major technological achievements during their formative years. They were able to identify four technology generations (using the translations of Docampo Rama, 2001): the *early technical generation* (born before 1938), the *household revolution generation* (born between 1939 and 1948), the *increasing technological household generation* (born between 1949 and 1963) and the *computer generation* (born in 1964 or later). According to their theory, different technology generations behave differently with respect to technology, due to different technology experiences in their formative period. As such, the concept of a *technology generation* might be a useful classification and explanans even for current older adults, struggling with technology that they did not grow up with. Docampo Rama (2001) breaks up the cohorts born in the second half of the 20th century into the *electro-mechanical generation* (born before 1960 and confronted in their formative years mainly with electro-mechanical interfaces on VCRs, televisions, radios etc.), and the *software generation* (born after 1960), which was heavily influenced by the transition to displays, touch buttons, wireless controls etc. in the eighties. The latter generation can be split further into the *display generation* (born between 1960 and 1970) and the *menu generation* (born in 1970 or later).

The concept of technology generations, however, is also problematic. First, the division into separate groups is largely based on theoretical considerations without a firm empirical basis. Many assumptions have to be taken, some of them rather arbitrary, such as technology membership (e.g. remote controls or teletext are assigned typical technologies of the *software generation*), technology diffusion (Docampo Rama uses the criterion of 20% of households to which a particular technology must have spread), or the rather broad formative period (children that experience a technology at 10 are clustered with adults experiencing technology at 25). Second, the granularity of the generations can be argued upon: is electro-mechanical technology a meaningful cluster, or would it be more telling to separate more fine-grained between telephone, radio, television or VCR generations? Third, it seems that technology evolution has accelerated to such a pace that technology generations might have become obsolete as soon as they are identified: does it really make a difference if a user belongs to the *display* or *menu* generation, when the interface under investigation uses touch & gesture? No generation differentiation has been proposed yet on

the basis of recent, but undoubtedly influential technological innovations, such as internet access, video games, mobile phones, social networks, mobile web or touch & gesture interfaces. Last but not least, even if valid and selective generations can be found, there still remains the intrinsic confoundation of generation and age effect in cross-sectional research, since both are measured by year of birth. In fact, the confoundation between age, cohort and period (known as the *identification problem* in cohort analysis), cannot be resolved statistically, but only estimated through laborious study designs (such as repeated cross-sectional studies), and analysis methods (for an extensive discussion of the problem and methods see Glenn, 2005).

2.1.3 Typologies

Another method to classify older adults that originates primarily from marketing research is the introduction of *typologies*. This method groups older people according to common values or lifestyles, in order to tailor products for a specific *type* of older user. Maier (2008), for example, suggested a division of the older customers into three major subgroups: the *traditional-passive*, the *culture-oriented active*, and the *adventure-oriented active*. Even though he uses this typology for senior friendly product design, those categories were mainly derived from the analysis of older people's spare time activities, and their usefulness to classify technology interaction has yet to be determined. Vandercammen (2007), in an interview study with over 1000 adults aged 50+, established a different typology with five categories: the *young* seniors, the *active* seniors, the *very old* seniors, the *peaceful* seniors, and the *hesitating* seniors. Such typologies are useful in determining market shares or product affinities, and for determining current trends, such as the growing number of active and hedonic individuals among the older adults (Petras, 2006). The existence of different categories illustrates the need not to speak simply of *the* older user, but instead to differentiate user types with varying backgrounds, attitudes, beliefs and values. At the same time, their existence (and moreover the lack of a commonly accepted typology in the scientific community) also emphasizes the extreme difficulty of putting this heterogeneous user group into any meaningful clusters. Too little empirical validation and measurement instruments exist up to now in order to use such a typology as a framework for studies in gerontechnology.

2.1.4 Terminology Used

Concluding, the concept of an *older user* is very hard to pin down. Neither does a commonly accepted age criterion exist, nor are there any established ways of subdividing the elderly population, being it according to age bands, generations or types. For the remainder of this work, we decided to adopt the viewpoint of Fisk et al.

(2004, p. 7) in their standard work *Designing for older adults*: “If forced to ‘give a number’ to the question of when a person is an ‘older adult’, we would most likely say that older adults are those individuals who are 60 years of age and older”. We also take up their basic classification into younger-old (60-75) and older-old (75+). Along with the literature, the term *older adult* is preferred to *old adult*. Even though there might be small differences in connotation, the terms *older adult*, *elderly*, or *aging adult* are used synonymously in the following, for stylistic reasons. We speak of *older user* whenever the older adult’s role in technology interaction is addressed.

For reasons discussed in Section 4.3, the focus of this research lies within the younger segment of the older adults. The empirical studies therefore concentrate on individuals aged between 60 and 75. This age band corresponds to the *younger-old* (Fisk, et al., 2004), but will be referred to as just *older adults*, for reasons of simplicity. It spans the silent generation and baby boom generation in Becker’s (2000) terminology. In terms of technology generations, the target group of this research belongs to the *electro-mechanical generation* (Docampo Rama, 2001), and includes members from the *early technical generation*, the *household revolution generation*, and the *increasing technological household generation* (Sackmann & Weymann, 1994).

2.2 Physiological and Psychological Changes with Age

Facing such big problems of framing what *older adults* refers to, the question arises whether it makes sense at all to think of the older adults as an entity, instead of just individuals. The group of older people is indeed highly heterogeneous, with each individual being shaped through sixty or more years of personal experiences. Despite these distinctive individual differences, older adults still have much in common in terms of the biological, psychological and social dimensions (Fisk, et al., 2004). These similarities can provide a starting ground to optimize product design with respect to the older adults.

A wealth of scientific literature already exists which investigates the changes in resources – both physiological and psychological – that accompany the normal aging process. These age-related changes can be clustered into alterations affecting sensation and perception, cognitive changes, e.g. in domains like memory, attention or intelligence, and adjustments in motor control. Excellent comprehensive overviews on these changes already exist, some of which even include design implications (e.g. Fisk, et al., 2004; Hawthorn, 2000; Martin & Kliegel, 2005; Mitzner, et al., 2006; Nichols, Rogers, & Fisk, 2006; Schieber, 2003). The following sections provide a brief overview of the various age-related resource changes, focusing on those that affect

the interaction with technology, since they are vital to the understanding of the present research.

2.2.1 Sensation and Perception

Several sensory and perceptual changes occur with age. Given that the most prominent modalities for interacting with technology (as well as with the rest of the world) are probably vision and audition, changes in these domains have significant effects on the way how older people handle technology.

Vision. Reduced vision capabilities are probably the most cited reason for problems: buttons on a remote control are too small to discern, the user manual is written in too small fonts, screen elements cannot be identified etc. Age is the single best predictor of visual limitations (Fisk, et al., 2004). However, age-related visual deficits are arising at multiple levels. Anatomical changes of the visual system include for example the reduction of pupil diameter in low light (Loewenfeld, 1979) and the opacification (loss of transparency) of the lens, resulting in an estimated 30% of light reaching the retina of a 60 year old as compared to a 10 year old (Weale, 1961). Visual acuity – the ability to resolve fine spatial detail – is deteriorated significantly, especially the ability to focus at near objects, and particularly in low luminance environments (Schieber, 2003). Furthermore, there is evidence of an age-related hypersensitivity to glare (e.g. Schieber & Kline, 1994) and reduced contrast sensitivity, beginning around 40 to 50 years (Owsley, Sekuler, & Siemsen, 1983). Older adults also suffer from a reduced peripheral visual acuity, reduced motion sensitivity and show more problems than younger observers to discriminate colors (Schieber, 2003). Taken together, these findings provide broad evidence that vision deteriorates significantly with age, and that deficits become more pronounced as the vision task becomes more challenging, for example with low contrast stimuli, poor illumination, peripheral or very near spatial position, or swiftly moving targets.

Hearing. The sense of hearing similarly reduces with age, leading to a significant increase in stimulus intensity required to detect a sound. This is especially pronounced for sounds at high frequencies (Fozard, 1990), and proceeds at a rate of approximately 1dB per year for adults beyond 55 (A. C. Davis, Ostri, & Parving, 1991). Numerous studies have further brought evidence that older people have greater difficulties to discriminate sounds of similar intensity and/or frequency, and also to properly localize sounds in space (details are given in Schieber, 2003). These phenomena contribute to poorer speech perception among older adults (e.g. Jerger, 1973). Challenging listening conditions, such as background noise or simultaneous speech from multiple agents, exacerbate these problems. Since technological systems that communicate to the user using speech are still an exception, the age-related

hearing deficits seem to be less crucial for technology interaction. However, for the design of systems that make use of acoustic warning signals, auditory icons or earcons, these findings require a careful selection of stimulus location, frequency and intensity.

Touch. Less research has been carried out on age-related changes and the perception of touch, even though touch remains an important feedback channel to control for the success of manual input in the field of HCI. Tactile perception, together with kinesthetic proprioception, will inform the user where fingers are located (e.g. on a keyboard, mouse, or touchscreen), and whether an action was successful (e.g. a button press or a tap). Wickremaratchi and Llewelyn (2006) review available evidence related to the study of aging and touch, and reveal similar deficits as were documented for the vision and hearing senses. They report findings on significantly increased detection thresholds for various tactile stimuli, which were most pronounced for vibrating stimuli. Besides tactile sensitivity, also tactile spatial acuity is decreased, so that for example at the finger, tactile acuity thresholds are on average 80% higher in older (65+) than in younger (18-20) adults. The authors further argue that impaired spatial acuity at the fingertips has an effect on hand motor function and may translate into greater difficulties with tasks requiring fine manipulation (Tremblay, Wong, Sanderson, & Côté, 2003). All these insights need to be carefully considered when it comes to the design of texture or pressure points of interface elements, button spacing, or vibrotactile feedback. The senses of smell and taste also show age-related changes, but they usually play no role in technology interaction.

2.2.2 Cognition

Given that cognitive processes naturally build on sensory input, it is not surprising to find vast evidence for changes in cognitive performance with age. Since these complex cognitive tasks are dependent on sensory and perceptual performance, it is not always easy to disentangle them (discussion of this problem and solutions to minimize the risk of confoundation are found in Scialfa, 2002).

Research on cognitive aging comprises the issues of learning, memory, attention, spatial cognition, reasoning, general intellectual properties such as the fluid and crystallized aspects of intelligence, language comprehension and production, and executive functioning, such as problem solving, decision making and coping with multiple task demands. A century of research in all these domains has left us with an abundance of studies, of which detailed reviews and summaries can be found for example in Birren and Schaie (2001), Schieber (2003), Fisk et al. (2004), Schaie (2004), Martin and Kliegel (2005), Mitzner, (2006), or Nichols (2006). A short overview of the most basic findings relevant in the context of HCI is given in the following.

Memory. Memory is commonly believed to decline with age. However, research has proven that memory is not a one-dimensional concept, but can be broken down into *working memory* (the ability keep information active in memory for a short period of time and integrate it with other memory and sensory information, term coined by Baddeley, 1992), and *long term memory* (the ability to permanently store knowledge), which again can be subdivided into *semantic* memory (the ability to store factual information), *episodic* memory (the ability to remember events in the past), *procedural* memory (the ability to implicitly apply previously learned skills), and *prospective* memory (the ability to remember tasks planned in the future). This segmentation is important, since not all memory mechanisms are equally affected by age.

For interface design, working memory is the most relevant concept, because during an interaction sequence, several chunks of information have to be maintained and integrated to successfully complete tasks, e.g. selecting the appropriate command in a hierarchical menu. Unfortunately, working memory is clearly affected by age. Schieber (2003) lists a series of studies that have shown large age-related declines in working memory *capacity*, meaning the amount of information chunks that can be kept concurrently in memory. Moreover, the capacity deficit increases with the complexity of mental operations to be performed (Salthouse, Mitchell, Skovronek, & Babcock, 1989). A simple and straightforward design implication following from these findings is to consequently reduce information load and in-memory-transformations, e.g. mental rotations. Concerning the different domains of long term memory, there is a general agreement that procedural and semantic memory remain relatively intact, but episodic memory shows an age-related decline (Mitzner, et al., 2006; Nichols, et al., 2006; Schieber, 2003). However, even for episodic memory, age deficits can be reduced (or even eliminated) if recognition procedures are used to retrieve information chunks from memory rather than free recall (Schieber, 2003).

Attention. Attention is still a rather fuzzy concept in cognitive psychology, but describes in general a cognitive mechanism (or set of mechanisms) that decides which pieces of information will be processed at all, with which priority, and which information is passed over to the next processing stage. Similar to the concept of memory, attention can be broken into several subsystems, which are affected by age to different extents. A popular conceptualization of attention comprises at least three separate mechanisms: *selective attention* (the ability to focus on certain objects while suppressing others, and purposely shift this focus), *divided attention* (the ability to attend several chunks of information simultaneously, sometimes also called attentional capacity), and *sustained attention* (the ability to remain focused on a task over an extended period of time) (for a detailed review see e.g. McDowd & Shaw, 2000). Several studies have demonstrated that older adults have increasing difficulties to select relevant information among distractors, which can be interpreted in

terms of a pronounced inefficiency to ignore irrelevant information among older adults (e.g. Rabbitt, 1965). Studies on age and divided attention, usually assessed through performance comparison of dual task to single task situations, have shown that older people are generally penalized in divided attention situations, and that their difficulties are once more most pronounced in increasingly complex situations (e.g. McDowd & Craik, 1988). Reports of age effects on sustained attention are not as abundant as the literature on selective and divided attention. A literature review by Giambra (1997) found equivocal evidence of a performance decline in sustained attention tasks, since older adults tend to miss rare events more often, independent of time and task. Taken together, there is broad evidence on age related deficits of attention, which have to be carefully considered when designing tasks and interfaces. In fact, attention has been suggested to be a key moderating variable predicting performance with an input device (Greenstein & Arnaut, 1988; cited from McLaughlin, Rogers, & Fisk, 2009)

Other cognitive mechanisms. Other aspects of cognition that have shown age-related changes include spatial cognition, language and general intellectual properties. Younger adults outperform older adults in tasks that require spatial transformation or navigation through space (Fisk, et al., 2004). Spatial abilities are relevant in the context of HCI for example in mobile phone menu navigation, where a performance decline with respect to age and spatial abilities has been demonstrated (Ziefle & Bay, 2006). Language and cognition are intrinsically related, and age effects on language production and comprehension are well documented (see Fisk, et al., 2004; Mitzner, et al., 2006; Nichols, et al., 2006). These will not be discussed here in greater detail, since they are less relevant to the design of interfaces that do not involve speech.

How all these different *cognitive mechanisms* relate and orchestrate to form *cognitive products* that manifest themselves in intelligent behavior is far from trivial, and a large variety of theories and models exist (see Schaie, 1996 for an overview). Several large scale longitudinal studies have investigated individuals on a broad battery of perceptual-motor and cognitive tests (e.g. Salthouse & Ferrer-Caja, 2003; Schaie, 1996; Schroeder & Salthouse, 2004). Common to these studies is the gradual decline on almost all constructs and, importantly, across the whole lifespan, and not merely from age 50 to 80 (Schroeder & Salthouse, 2004). Correlation analysis and structural modeling reveal what has been already pointed out above: Many of these different variables are significantly correlated and share explanatory power, giving rise to a common factor in the structural equation models that seems to be responsible for many of the observed effects (Salthouse & Ferrer-Caja, 2003; Schroeder & Salthouse, 2004). Salthouse and Ferrer-Caja (2003, p. 107) suggest at least three distinct mechanisms to account for the observations: “(a) one to account for common age-related

effects across reasoning, memory, and spatial abilities; (b) one to account for additional age-related effects on speed; and (c) another to account for additional age-related effects on memory”.

The fact that many degradations in cognitive performance, especially in what has been labeled *fluid cognition* can be attributed to a general slowing of processing speed had been postulated in the “Processing-Speed Theory” before (Salthouse, 1996) and is in line with these findings. The division of a general intelligence factor into *fluid intelligence* (incorporating aspects like learning, logical reasoning and problem solving) and *crystallized intelligence* (including skills, knowledge and experience, Cattell, 1963; Horn & Cattell, 1967), might indeed be a useful conceptual framework to account for age-related differences in cognitive performance, since there is consensus that fluid parts of intelligence show a greater age-related decline, while crystallized aspects are largely preserved (e.g. Horn & Cattell, 1967). Building on these results, Baltes, Staudinger and Lindenberger (1999) have developed a model of intellectual development across the lifespan that distinguishes two main categories: the *mechanics* and the *pragmatics* of cognition. *Mechanics* comprises the cognitive aspects involved in information processing that are universal, genetically predisposed and correspond to the fluid aspects of intelligence. *Pragmatics*, on the contrary, is experience based, culture-dependent, and corresponds to the crystallized aspects of intelligence. A large body of empirical evidence confirms the different trajectories of these two components across lifespan: “Abilities that critically involve mechanics, such as reasoning, spatial orientation or perceptual speed, generally show monotonic and roughly linear decline during adulthood, with some further acceleration of decline in very old age. In contrast, more pragmatic abilities such as verbal knowledge (e.g. semantic memory) and certain facets of numerical ability, have weak, and sometimes positive, age relations up to the sixth or seventh decade of life, and start to decline only in very old age” (Baltes, et al., 1999, p. 487).

2.2.3 Motor Control

Finally, also the control of movements worsens with age. Older adults are generally slower in movement execution than younger adults and their movements are less precise (Fisk, et al., 2004). Changes in older adults related to movement control include reduced movement speed, movement composition differences, increased variability of movements, reduced force control, and coordination difficulties (for a review, see Ketcham & Stelmach, 2004). It is easily conceived that for example the control of a computer mouse has considerable demands on trajectory control, end-point control, force regulation and the coordination of movement components, all of which show some age-related decline (Ketcham & Stelmach, 2004). Several studies have indeed demonstrated that cursor control with a mouse becomes more difficult

for older people, evidenced for example through longer movement times, more sub-movements and a greater number of errors (Charness, et al., 1995; M. W. Smith, Sharit, & Czaja, 1999). Particular problems of older users related to different input devices will be discussed in greater detail in Section 2.5.2., and more details on motor control deficits and their implications for gesture control are given in Section 5.1.

2.3 Attitudes and Anxiety Towards Technology

The previous section has given a brief overview of the sensorial, perceptual, cognitive, and motor control changes that the normal aging process brings about, and outlined the implication for technology interaction. Evidence is brought forward mostly through interaction performance measures: the time an older user needs to complete a task, the number of information pieces that can be handled, the number of errors that occur during interaction, etc. While there is no denying the fact that older adults have greater difficulties *while using* a device (and many examples of this are given in the next section), the factors that drive – or hinder – older adults *to use* technology in the first place also need to be considered. In fact, older people are often believed to have a disinclination towards technology, to be afraid of using unknown products, and generally hesitant to adopt new technologies.

First, it is important to note that *attitudes* and *anxiety* towards technology should not be confused, even though these two concepts share some variance (e.g. Igbaria & Chakrabarti, 1990 found a correlation of 0.32 between a measure of computer attitude and a measure of computer anxiety). *Attitude* is a concept from social psychology, which refers to one's beliefs, feelings and behaviors toward an object (also a person or event), e.g. a computer or other piece of technology. *Anxiety*, on the other hand, is a clinical term, referring to a negative and/or stressful emotional state when thinking about an object (e.g. computer) or using it (Kelley & Charness, 1995). However, anxiety is also sometimes used as a subscale in computer attitude instruments (Gardner, Discenza, & Dukes, 1993).

Attitudes. Attitudes tend to guide behavior (Regan & Fazio, 1977), and thus it has been suggested that older adults are reluctant to use computers because of poor attitudes towards them. The major part of research, however, has found no difference in attitudes towards computers between younger and older adults (for a review, see Kelley & Charness, 1995). Jay and Willis (1992) assessed older users' attitudes towards computers using a multidimensional questionnaire, including comfort, efficacy, gender equality, control, dehumanization, interest and utility. They found that older adults' computer attitudes are not static but can be altered and that immediate

computer experience is an effective means of change. Czaja and Sharit (1998), using the same instrument, confirmed these findings by showing that experience with computers could increase participants' feelings of comfort with technology, their competence and perceived utility. More importantly, the multidimensional account of attitudes towards computers allowed them also to single out dimensions where older adults showed less positive attitudes than the younger comparison group: they reported less comfort and less competence with computers, perceived computers as more dehumanizing and felt they had less control over the computer, even if controlled for computer experience. Taken together, it seems that, even though older adults might be less positive on certain dimensions of technology related attitudes, these can be modified by experience and learning, and no global effect of attitude can account for a poor adoption of technology by older adults.

Anxiety. Computer *anxiety* has been linked to a perceived loss of control, unfamiliar computer terms, and a fear of negative evaluation by self and others (Meier, 1985). Laguna and Babcock (1997) investigated the relationship between age, computer anxiety and performance on a computer-based cognitive task. They found that older adults (55-82 years) had a significantly higher computer anxiety than the younger group (18-27). However, anxiety did only show an influence on performance with regard to decision time, and not for task acuity. Moreover, other studies (e.g. Charness, Schumann, & Boritz, 1992) did not find significant age differences in computer anxiety. Kelley and Charness (1995, p. 114) state in their review that "the influence of anxiety or negative attitudes and age on computer performance is of limited generalizability and has yielded mixed results", which can be explained to some extent also by the fuzziness of the computer anxiety construct and the lack of standardized measures, rendering comparison across empirical studies very difficult (Dyck, Gee, & Al-Awar Smither, 1998).

2.4 Technology Acceptance and Adoption

In the sections above older adults have been characterized by the changes in perceptual, cognitive and motor resources as well as their attitudes and anxiety towards technology. How these factors orchestrate together towards a generally low technology *acceptance* and *adoption* among the elderly, and what is their relative impact, is largely still unknown. Some models exist that shed more light in this complex interplay of factors.

Technology *acceptance* and *adoption* have become key concepts in a broad range of research fields, spreading from psychology and sociology to ergonomics and market-

ing (Arning & Ziefle, 2007; Renaud & van Biljon, 2008). While often mentioned in the same breath, these terms have a slightly different focus. Technology *adoption* is preferred in sociological studies and describes a macro-level *process*, including everything from a purchase decision, to actually using a system and accepting or rejecting it (Gelderblom, Dyk, & van Biljon, 2010). Technology *acceptance* is described as “the approval, favorable reception and ongoing use of newly introduced devices or systems” (Arning & Ziefle, 2007, p. 2905), and focuses more on the *factors* contributing to the approval or disapproval of a system. Since both concepts - though highly similar - have different perspectives, there is value in combining them (Y. S. Lee, 2007). Along with Gelderblom et al. (2010), we perceive acceptance as a necessary constituent of the adoption process.

Czaja et al. Czaja and colleagues (2006) postulated an influence of age, education, computer efficacy, computer anxiety as well as fluid and crystallized intelligence on technology adoption. Their results confirmed that a digital divide still exists: older adults have less technology experience in general, and also less experience with computers or the World Wide Web in particular. Concerning the relationship of the investigated factors, they found that especially computer anxiety and intelligence were important predictors of technology use, while the relationship between age and technology adoption was mediated by cognitive abilities, computer self-efficacy and computer anxiety. Furthermore, the authors report a strong independent effect of age on technology acceptance that is not accounted for by attitudes and cognitive abilities, indicating that further age-related factors not captured in this study also play a role in the decision to adopt a new technology.

Renaud and Biljon. Another approach to a holistic understanding of the relationship between older adults and their use of technology is the Senior Technology Acceptance & Adoption Model (STAM) proposed by Renaud and Biljon (2008), which is based on the hugely influential *Technology Acceptance Model* (TAM, F. D. Davis, 1985, 1989) and its descendants TAM2 (Venkatesh & Davis, 2000), UTAUT (*Unified Theory of Acceptance and Use of Technology*, Venkatesh, Morris, Davis, & Davis, 2003), and MOPTAM (*Mobile Phone Technology Acceptance Model*, van Biljon & Kotzé, 2007; van Biljon & Kotzé, 2008). Central in Davis’ original model were two main determinants affecting the *behavioral intention to use* (BI) a certain technology: *perceived usefulness* (PU), and *perceived ease of use* (PEOU). Both factors are influenced by a range of external variables, such as for example training with a system and documentation of a system (F. D. Davis, 1989). However, even in the revised TAM, as pointed out by Arning and Ziefle (2007, p. 2906), “the impact of user characteristics like demographic variables (age, gender, expertise) or cognitive (abilities) and attitudinal (self-efficacy) concepts on perceived ease of use and usefulness ... is not adequately regarded yet”.

STAM differs from previous TAM-inspired models by including a temporal dimension, consisting of an *objectification phase*, *incorporation phase* (experimentation and exploration, influenced by PU and PEOU), and *conversion phase*, which were borrowed from the *domestication of technology* theory by Silverstone and Haddon (1996). In the last phase, after actual usage, older users would either accept the technology (conversion), or reject it (non-conversion). An important difference to previous models is also that age is not considered to be a mediator of the *intention to use*, in accordance with results by Phang et al. (2006). The authors argue that older adults do *intend* to use technology; however, it's the reduced ease of use which they experience that hampers the actual usage of a product. The major contribution of STAM is relating the acceptance factors to adoption stages, and by this providing an explanation why many elderly adults never reach the final adoption phase and never fully accept the technology. But STAM is not without criticism. An evaluation study by Gelderblom et al. (2010) revealed several shortcomings of STAM: The dichotomy of accepting/rejecting does not map into reality very well, since many older people do not fully reject or accept a technology, but rather judge it somewhere in between. Moreover, users may never actually move on to the conversion phase, but just continue to use a technology without appraisal. Finally, the authors also found that actual use is not always preceded by an experimentation and exploration phase.

Arning and Ziefle. Another model trying to predict older users' acceptance of technology was suggested by Arning and Ziefle (2007). At the heart of their model were the original constructs PU and PEOU from Davis' model, but they integrated *subjective technical confidence* (STC) as an additional factor to influence PEOU. Most interestingly, they place performance (= the operationalized actual use of a system) as a predictor for PEOU, while the original TAM sees PEOU as predictor for BI, from which actual use follows. Indeed, regression analyses showed that performance was found to be the main predictor of PEOU (Arning & Ziefle, 2007). STC, in this model, did not explain significant parts of variance in PEOU, whereas Wilkowska and Ziefle (2009), with a highly similar model, found the contrary.

Conci et al. Conci, Pianesi and Zancanaro (2009) recently suggested yet another model to account for mobile phone adoption among older people, which extend TAM by including further antecedents of PEOU, notably *enjoyment*, *self-actualization*, and *perceived safety*. Their main contribution is the distinction between intrinsic motivational factors to use a mobile phone and extrinsic motivators. Their results show that both *enjoyment* and *self-actualization* have significant influence on PU and PEOU. Furthermore, also *perceived safety* could be confirmed as significant influence to mobile phone acceptance by increasing its utilitarian value. In their model, the strongest predictor for using a mobile phone was its perceived usefulness for the older adult.

Conclusion. The review of the above models leads to several conclusions. First, there is value in combining the research on acceptance and adoption of technology, since the concepts are so closely related and share a lot of variance.

Second, the literature on technology acceptance and adoption is far from converging. On the contrary, it seems to be the case that new models or model extensions are sprouting for every context or application, such as work (Venkatesh, et al., 2003), mobile phones (Conci, et al., 2009; van Biljon & Kotzé, 2007), PDAs (Wilkowska & Ziefle, 2009), eHealth (Arning & Ziefle, 2009) or eGovernment (Phang, et al., 2006). Bagozzi, in a critical assessment of the development of the TAM model and its various extensions, expresses quite drastically his opinion that “the field is at a threshold of crisis, if not chaos, in regard to explaining technology acceptance” (Bagozzi, 2007, p. 244). He rightfully acknowledges that “the main strength of TAM is its parsimony”, but at the same time that “Parsimony has also been an Achilles’ heel for TAM” (Bagozzi, 2007, p. 244). A whole group of TAM descendants have added many additional factors to adapt to specific contexts, potentially increasing explanatory power, but at the same time limiting its generalizability and, maybe more importantly, rendering cross-model comparisons increasingly difficult.

Third, concerning computer efficacy (CE) and computer anxiety (CA), no consistent picture emerges: while Venkatesh et al. (2003) report no significant effects of CE and CA on BI, Czaja et al. (2006) found a direct effect of CA, and an indirect effect of CE on computer use.

Fourth, except for the model by Czaja, which included crystallized and fluid intelligence as predictors of technology use, the other models show a striking disregard of the vast body of literature on age-related changes in abilities. While predictors related to the specific technology (such as support, perceived safety etc.), related to the use context (social influence, facilitating conditions), or the interaction itself (PEOU, PU) are included, factors like memory, perceptual speed, psychomotor speed etc. are not accounted for.

Lastly, the models focus mainly on *intention* to use, rather than the actual use. Actual use is only *implied* by intention to use, a linkage based on Fishbein and Ajzen’s (1975) Theory of Reasoned Action (TRA), and has been described as “intention-behavior gap” by Bagozzi: “From my point of view, the intention-behavior linkage is probably the most uncritically accepted assumption in social science research in general and IS [=information systems] research in particular” (Bagozzi, 2007, p. 245). One issue with this linkage is the fact that usage is often not so much an end in itself, but usually means for a more fundamental goal, something not accounted for in the TAM models. Another issue is the temporal dynamics: between intention and actual use can be many intervening steps and unforeseen events or obstacles, upon which intentions get re-evaluated and decisions to use potentially changed (Bagozzi, 2007).

By reason of the issues raised above, such as the lack of convergence, the lack of consistency, the lack of inclusion of user-specific abilities, the lack of predictive power for actual use, and few available validation studies on the age-specific models, there is no particular model singled out to guide this research. Rather, this research takes an applied perspective, focusing on the actual use of a certain technology by elderly people. We further shed some light onto the relationship between older people's abilities and the demands that a gesture-based interaction brings about.

2.5 Problems during Technology Interaction

The previous sections gave a rough overview on the resource changes with age, on attitudes and anxiety towards technology among older adults, and presented many more factors that need to be considered for technology acceptance. In contrast to the well-documented age-related decline in resources, older adults are generally still very proficient at certain tasks (which could also include the use of technology), because they employ particular strategies, such as *selection*, *optimization* and *compensation* (c.f. SOC model by Baltes & Baltes, 1990). That is, by concentrating on domains of high priority, by enriching and augmenting their general reserves, and by switching to new strategies and technologies (e.g. using new mnemonic strategies, hearing aids etc.), older adults arrive at a "reduced and transformed but effective life" (Baltes & Baltes, 1990, p. 22). With regard to the usage of technology, this means that hypothesized problems related to their abilities (see previous sections), might not become effective in actual use, and that the use of certain strategies or additional aids may change their willingness to use a technology as predicted by various models. It is therefore of interest to scour the literature for *actual problems* encountered by older users while using technology. Since the focus of this research lies on mobile devices and a novel input technique, the listing of interaction problems will concentrate on these aspects.

2.5.1 Mobile Interaction of Older Adults

Research on mobile interaction among older adults first needs to recognize the different usage patterns that older people exhibit in using their mobile devices. Older people use mobile phones quite differently from younger users in terms of frequency, intensity and use context. Kurniawan, for example, found that older people typically do not use their mobile phones for casual conversations, but restrict their use most often to emergencies (Kurniawan, 2006, 2008; Kurniawan, Mahmud, & Nugroho, 2006). In fact, all participants cited the increased feeling of safety and

security as their prime reason for carrying a mobile phone, and could not think of any other real benefits of carrying a mobile phone around (Kurniawan, 2006; Kurniawan, et al., 2006). Furthermore, they showed to be cost sensitive, and only resorted to call a person via mobile phone if that was the cheapest way to contact. The list of potential contacts usually proved to be extremely limited. Older users were calling mostly family members, and only rarely their friends. Of interest is also their very limited use of short messaging. As one reason (apart from obvious problems related to text reading and text entry, see **Table A.1-1**) Kurniawan (2006) cites the older adults' belief that SMS would ruin people's literacy. In fact, older people can generally be described as passive users, since they mostly use their mobile phones to be contacted themselves (either through calls or text messages), rather than actively make contact. The fact that most functions beyond making and receiving voice calls are so rarely used might be explained by the older people's view on the mobile phone: they consider it barely a modern version of a landline phone (Kurniawan, 2008).

Recent research has suggested, however, that also among older users, the view of the mobile phone seems to change from a supplement of other means of communication, only resorted to if other options are not available, to an important companion in their daily routines. For example, the availability of functions like SMS, MP3 player, camera function etc. has been stated as third-important usage motive in focus group sessions by Arning and colleagues (2010). The two most important usage motives were facilitation of daily life activities and the improved availability of oneself and others. The safety motive only ranked 10th. Even though the sample in Kurniawan's and Arning's study cannot really be compared because of large differences in sample size, cultural background, gender and educational composition, it might also be the huge success of smart phones that occurred between 2006 and 2010 that led to a change of the public's perception of the role of a mobile phone. Very interesting are also the most prominent barriers to use mobile technology stated by older users. Above everything else, older users are still afraid of health damage and radiation from the mobile phones (Arning, et al., 2010). Other barriers frequently mentioned include technical unreliability, and issues of privacy and surveillance. It has to be noted, that the specific motivators and barriers are also subject to change in differing use contexts, e.g. personal communication as opposed to healthcare technology (Arning, et al., 2010).

Despite rising adoption rates of mobile phones among older people in recent years (e.g. Marketing Charts, 2010; A. Smith, 2011), there is still a large gap between mobile phone usage between American adults aged 18-65 (on average 89% own a cell phone), and those aged 66 and above (on average 58% own a cell phone) (Zickuhr, 2010). The single most often reported issue concerning the use of mobile technology

by elderly people remains the difficulty to use it. For example, in a study by Osman and Maguire, older adults were asked to rank desired features for a mobile phone in three categories, related to usability/ergonomic features, phone functions and advanced services (Osman, Maguire, & Tarkiainen, 2003). All features related to usability/ergonomic issues were ranked as the most desired features. In a similar study, where older adults were asked to rate the importance of several product attributes of a mobile device, they rated ergonomics and ease of use as the most important factors (Pohlmeyer, Machens, & Blessing, 2010). Furthermore, the authors could show that these factors were considered more important by the older adults than by a younger comparison group.

The problems elderly encounter when trying to use current mobile phones are related to specific interface elements (e.g. buttons or icons), specific tasks (e.g. text entry) or specific functions (e.g. camera), but can also relate on a more general level to incorrect mental models of the menu hierarchy and other conceptual issues. The long list of problematic issues regarding the use of mobile devices by older adults (see Appendix, **Table A.1-1**) shows in the first place that the design of these devices is in many ways not yet tailored to their needs and abilities. The problem is still of high relevance, since the cited evidence (cf. **Table A.1-1**) spans only last six years. It further shows that substantial problems prevail, despite rising adoption rates, despite older adults' willingness to use technology (Czaja & Lee, 2007) and despite compensatory strategies (Baltes & Baltes, 1990). It is only reasonable to assume that the daily encounter of such interaction problems leads older people to name usability and ergonomic issues first when asked about most crucial barriers, and it also justifies the prominent position of perceived ease of use (PEOU) in the models describing technology adoption among elderly. "Generally, the existing literature shows that, although older adults are willing to use technology, many report usability problems with existing systems and these problems may in part be due to the cognitive and perceptual demands placed on the user" (Czaja & Lee, 2007, p. 341).

Several researchers have already compiled lists of design features desired by older people, and inferred design guidelines for more senior friendly mobile phones (e.g. Asano, et al., 2007; Czaja & Lee, 2007; Irie, Matsunaga, & Nagano, 2005; Kim, et al., 2007; Kurniawan, 2008; Lin, Hsieh, & Shiang, 2009; Osman, et al., 2003; Pattison & Stedmon, 2006; Ziefle & Bay, 2004). Many of these guidelines are not restricted to mobile phones, but in accordance with general design recommendations for older adults, such as found for example in Fisk and colleagues (2004) or Nichols and colleagues (2006). However, designing for an older generation is not a straightforward undertaking, simply ticking off the recommendations found in the several design guidelines. First, even though of highest importance, ease of use and ergonomics are only a part of a complex interplay of product attributes, among them, for example,

aesthetics, quality, functionality, emotional appeal, cost etc., which only in their conjunction yield truly desirable products for older users (Pohlmeyer, Blessing, Wandke, & Maue, 2009). Second, desired features are often contradictory, so that one feature cannot be optimized without harming another feature. For example, older users desire a phone which is thin and lightweight (implicating small dimensions), and at the same time they want a phone with a big screen and big buttons (implicating large dimensions) (Kim, et al., 2007). Obviously, compromises need to be made that potentially do not match the specific needs and capabilities of certain users.

2.5.2 Benefits and Drawbacks of Different Input Types

Since the touch and gesture interaction under investigation in this research is essentially a new input paradigm, it is worthwhile scrutinizing the literature on input devices for older users. Several studies have investigated how different input devices might facilitate or hinder interaction performance for the older user, mostly in the context of computer interaction (for an overview of different input devices for older users, see Taveira & Choi, 2009). Insights from these studies could allow anchoring gesture interaction in the space of different input styles and devices, and thus fathoming its potential benefits.

Rogers, Fisk, McLaughlin and Pak (2005) draw some important conclusions from their research: First, different input devices affect computer task performance. Second, the optimal input device to facilitate interaction is dependent on the age of the user. But also third, the optimal input device is further dependent on the type of task, thus it is inappropriate to suggest one input device as always the best choice. The authors make an important distinction between *direct devices* and *indirect devices*. A *direct input device*, such as a touch screen (including multi-touch & gesture), stylus or light pen, translates the movements of a user directly to the manipulation of objects in the interface (e.g. the movement trajectory of the finger is isomorphic with the trajectory of an object dragged on the screen). *Indirect input devices*, such as a computer mouse, a mini-joystick, a trackball, or a scroll wheel, translate the user input from input to output domain, usually using gain and sometimes transforming the movement plane (e.g. movement of the mouse *towards* the user over a *small distance* results in a movement of the cursor *downwards* on the screen over a *larger distance*). Rogers and colleagues (2005) give an overview of the general advantages and disadvantages related to the two interface groups, which are summarized in **Table 2.1**.

Table 2-1: Comparison of input devices (adopted from Rogers, et al., 2005)

Device	Advantages	Disadvantages
Direct Devices		
Examples: touch screen, light pen, voice recognition	Direct hand-eye coordination	Arm fatigue
	No need to memorize commands	Limited resolution
	Minimal training	Difficulty with precision
	High user acceptance	Slow entry
	Requires less space	Finger or arm may obscure screen
	Long, ballistic movements accomplished quickly	Inadvertent activation
	Better for pointing tasks	No inherent feedback
Indirect Devices		
Examples: Rotary encoder, mouse, joystick, trackball	Can adjust control-display ratio	Requires translation between rotary and linear movement
	More precise	
	Gives tactile feedback	Requires translation between hand and screen
	Experienced users prefer it for long periods of use	Requires learning time
		Movement time between controls is lengthy

Comparative studies comparing different input devices for older adults find advantages of direct input devices over indirect ones (Armbruster, Sutter, & Ziefle, 2007; Chaparro, Bohan, Scarlett, Fernandez, & Choi, 1999; Charness, et al., 1995; Charness, Holley, Feddon, & Jastrzembski, 2004). Other comparative studies point rather at a task-dependency regarding which input device works best for older adults (Jastrzembski, et al., 2005; McLaughlin, et al., 2003; McLaughlin, et al., 2009; Rogers, et al., 2005). In a review study by Sutter and Müsseler (2007), input devices for older adults were not compared along the direct - indirect dichotomy, but along a continuum of sensorimotor transformation, ranging from very easy (pen) to easy (mouse, touchpad) to difficult (trackpoint). Even though their older sample was relatively young ($M=51$ years), they could still document that age differences were largest for the trackpoint. The age difference dropped as the sensorimotor transformation was reduced, until it disappeared for the pen. The authors speculate that these effects may become more pronounced at higher age.

Only few studies compare touchscreens with other input devices. Murata and Iwase (2005) found evidence of a significant age-related slowing in pointing tasks for mouse interaction, while touchscreen interaction showed no age-related decline in

performance. These findings could be replicated in a similar study by Murata and Takahashi (2008).

A comprehensive comparison of four input devices for older users (touchscreen, trackball, mouse and touch pad) found strong effects of experience (Wood, Willoughby, Rushing, Bechtel, & Gilbert, 2005). Older users who had frequent daily experience with computers performed best with the mouse – the device they were most familiar with. Novice users, on the contrary, fared better with the trackball, which has lesser demands on fine motor coordination. The touch-based devices were found to fall behind on performance measures, and the main reasons for this were identified as (1) obstruction of the view by the arm, and (2), the need to exert sustained pressure.

Notably, all comparative studies involving older users found to date only investigated computer input methods, and not input methods on small screen. The problems with touchscreen input (obstructed view & sustained pressure) might be substantially reduced with more recent capacitive touchscreens, and with only one finger obstructing the view for handheld devices.

Summarizing the various studies, it becomes clear that no particular device can be regarded as the generally best suited input method for older users. Rather, “it is critical, especially for older adult users, that there is a match between the task demands and the device characteristics. The performance of older adults is influenced more by a task/device mismatch and by demands of their attention” (Rogers, O'Brien, & McLaughlin, 2006, p. 5). Nevertheless, there seems to be an effect of sensorimotor transformation, rendering devices with higher transformation demands increasingly difficult to use especially for older users. This accounts for the tendency of direct interaction devices being easier to use by the elderly than indirect interaction devices. However, it seems that touchscreens, being a direct interaction device, are more advantageous for users equally inexperienced with alternative input modes, and only when the disadvantages (e.g. obstruction, sustained force, lower precision) are properly addressed. Research on the aptitude of current touchscreens in mobile devices, which have reduced size, but high resolution, high responsiveness and often multi-touch capabilities, is still scarce.

2.6 Designing for Older Adults

The previous sections have demonstrated that older users still face a multitude of problems during the interaction with technology. It has also become clear that many of the problems are specific to the older user group, or at least get more pronounced.

What can be done to improve this situation and grant older adults a better access to technology?

In our view, there are different strategies that complement each other to accomplish this goal: *Motivating* the older users to bridge their initial disinclination towards technology (e.g. Melenhorst, Rogers, & Bouwhuis, 2006), *training* older users with specific adaptive training concepts (e.g. Bruder, 2008), or *developing devices* tailored to the older user's needs and abilities. While the best results for improving the use of technology by elderly users are surely achieved by a combination of these strategies, our research is centered on the problem of interface design for the older user. Finding solutions to this problem is a tightrope walk between two pitfalls: On one hand, if technology is designed with a focus on specific user groups (e.g. *Design for Diversity*) (Gregor & Newell, 2001), such as the older users' characteristics, it runs the risk of stigmatization and reduced acceptance, in particular when the focus is on the deficiencies of older users. On the other hand, following strictly a philosophy of *Universal Design* (UD, Mace, Hardie, & Place, 1991; Story, 2007), *Transgenerational Design* (Pirkle, 1994) or *Inclusive Design* (ID, Keates & Clarkson, 2003; Persad, Langdon, & Clarkson, 2006) might not properly address the specific problems of the elderly, or lead to design compromises which are favored neither by the elderly nor by younger generations.

For example, this approach has resulted in senior-specific mobile phones with extremely large buttons, a large display, and a loud speaker, such as "Katharina das Große" (fitage GmbH, 2006). While critically acclaimed and commercially rather successful, the product targets at a very specific and limited user group and can also be criticized: First, the design is centered on users' disabilities instead of their abilities. Such products are often regarded as "hospital technology" (Dienel, 2007). Older users might be reluctant to buy such technology if it is only available in specialized stores (e.g. medical supply stores), or might be reluctant to use it in public if the design transcends openly that it was developed for impaired users. Goebel (2007) argues for the need of a subtle balance between ergonomic requirements on one side and social categorization on the other. Second, the concentration on certain aspects (e.g. extremely large interface elements) can come at the cost of other product attributes such as aesthetics, which has been shown to be relevant to older users just as well (Pohlmeyer, et al., 2009). Lastly, in being very specific in their design and function, such products intentionally dismiss certain functions that are common in comparable products (e.g. short messaging and alarm clock functionality are only optionally available in "Katharina das Große"), but in doing so, they exclude a large group of users that want these features.

The solution can only be a carefully balanced compromise, which is neither truly universal, nor accounting for each individual difference in abilities. Thorough re-

search is needed to separate design features that all age groups can profit from, from those that might bring a benefit to some, but are detrimental to others. For example, Schneider and colleagues (2008) concluded in their study of a project management software, that some changes in layout lead to better performances in all three age groups, whereas an increase in font might bring reading advantages for the older group at the cost of disadvantages for the younger group.

We see a huge potential in adaptable products. Particularly in the field of electronic devices, is it possible to adapt certain aspects of the software or firmware to the needs of particular users (e.g. different profiles on a device), while at the same time the unchangeable hardware features can be designed according to ID or UD philosophies. Other researchers have also advocated this type of profile-based solutions (e.g. Stone, 2008). Olwal, Lachanas and Zachrouli (2011), for example, presented a prototype framework which allows to customize interfaces to the needs of older adults, by decoupling the software UI from the phone's physical form factor. This approach – one might call it “Adaptable Design” – is especially promising for devices that rely heavily on touchscreen input. In those systems, the physical limitations of the system play only a minor role, since hardware keys are largely removed. These devices offer for the first time the chance to adapt not only the graphical output on the screen (e.g. adjusting font sizes, icon density etc.), but also input related parameters, like the size of touch-sensitive area, the sensitivity of touch recognition, or the alphabet of interaction gestures. General purpose devices could thus, in principal, be universally designed for many basic features, but then adjusted for diversity regarding more specific features.

2.7 Chapter Summary

This chapter has started out by defining the “older user” and discussed several approaches to divide the older population into meaningful clusters. Since definitions and classification schemes differ greatly, older users were pragmatically defined as those individuals aged 60 and beyond for the scope of this thesis.

The older adults were characterized with respect to age-related changes in sensory, cognitive and motor abilities, and these abilities were linked to a huge variety of problems that older users are still facing when interacting with technology. Three different aspects of technology interaction were elaborated in greater detail.

First, a review of models striving to explain technology acceptance and adoption among older users revealed a lack of convergence, consistency, and a disregard of the actual abilities and limitations of elderly users. None of these models was select-

ed to provide the framework for the current research. Rather, an applied perspective was taken that focuses on actual use instead of the mere intention to use.

Second, a review of studies dealing with older adults and mobile phone usage substantiated the claim that with current devices, there is still a gap between technology design and the needs of older users. A list with various sources of interaction problems was presented (**Table A.1-1**). The focus on mobile devices within the scope of this research is therefore justified.

Third, since the focus of this work is particularly on touch and gesture technology, a literature review of different input technologies with respect to their aptness for elderly users was given. A general benefit of direct input devices with little sensorimotor transformation (such as touch and gesture) could be shown. However, it has also become clear that not a single input mechanism is always best, as there is a strong task dependency.

Finally, different design philosophies that cover older and disabled users were presented. Although labeled differently, they are similar in their core, and ultimately all have to find compromises between specific but exclusive, and inclusive yet generic designs. It was argued that adaptive products could be a solution to this predicament, and that touch and gesture operated devices have particularly great potential in this respect, since so many interface parameters could be adapted to the needs of older adults.

3 Gesture Control

Since the gaming console Wii was released by Nintendo in late 2006, and Apple's iPhone started to revolutionize the market for mobile phones half a year later, gesture interaction has taken the transition from prototype applications or niche products to a widely accepted way of controlling everyday technology. Some researchers even speak of a "new era of interaction design", boldly claiming that "tap is the new click" (Saffer, 2008, p. 3).

The effects of both the Wii console and the iPhone on their respective markets have been huge. The Wii became Nintendo's best selling (non-portable) console ever (Nintendo Co. Ltd., 2011), outselling its direct competitors – Sony PlayStation 3 and Microsoft Xbox 360 – by one third, despite its significantly lower graphics and computing performance. What set it apart was the innovative way of controlling the system: three-dimensional movement patterns carried out in midair can be captured by the system and translated into game commands. A similar success story has been written by the iPhone, which has become one of the bestselling mobile phones ever produced, with more than 100 million units sold, and leading Apple from a newcomer in the mobile phone market in 2007 to the world's largest handset vendor in 2011 (Mawston, 2011). Reasons for this huge success can be found on various levels, but prominently among them is surely the innovative interaction concept, which brought multi-touch control for the first time to a mass consumer product. Never before did users navigate through lists and libraries so effortlessly with the tip of their fingers, did they zoom in and out of pictures, rotate maps, adjust volumes etc. just by the movement of their fingers on the screen. In recent years, this style of interaction has become the de-facto standard for the vast majority of mobile devices.

This new interaction style, in which movement patterns carry meaning and can be interpreted by a technical device has been termed "gesture control", "gestural interaction", "gesture-based interaction" or similarly. It is often associated with a certain "intuitiveness" or "naturalness" of interaction, and attributed to a particular ease of use. In Chapter 4, the prospects and potentials of this input technology with regard to the older users will be discussed in detail. Before, however, a clarification of the terminology is required. The following sections address the questions: What is meant by "gesture" in a human communication context and in a human-machine interaction context? How can gesture systems be described and classified? Which technical limitations do they imply?

3.1 What Is a „Gesture“?

The interaction required for the Wii is fundamentally different from the interaction required by an iPhone. Apple describes the movement patterns the user has to “draw” on the multi-touch screen as *gestures* (Apple Inc., 2011a, 2011b), and so does Microsoft for their Windows Mobile 7 control functions (Microsoft, 2011). However, Nintendo prefers to speak of “motion control” instead of gesture for their Wii controller (Nintendo, 2011), despite the fact that its three dimensional movements seem closer to natural gestures that occur in human-human communication, than the movement of fingers on a touch-enabled surface. But what exactly defines a gesture? Is the swinging of a tennis racket a *gesture* – or does the fact that the tennis rack is virtual (i.e. Wii control) make it one? Can a simple tap on a touchscreen already be regarded as a gesture, and if so, what sets it apart from a keystroke? Are the gestures used in human-computer interaction fundamentally different from the ones in human-human interaction?

3.1.1 Gestures in Human-Human Communication

Gestures in human communication are generally regarded as an effective means of nonverbal communication. They are used all the time, consciously and unconsciously, in the absence of verbal explanations or simultaneously to speech. While people usually have an intuitive understanding that forming a ring with thumb and index finger – signaling *OK*, *zero*, *money*, or other meanings depending on the culture – can be regarded a gesture, it is not clear where to draw the line between other types of hand or finger actions. Certainly, “... for an action to be treated as a ‘gesture’ it must have features which make it stand out as such” (Kendon, 1986, p. 26). What these features exactly are has been debated at length in linguistic and psychological research.

The first systematic investigation of gesture is probably David Efron’s seminal work from the 1930s. His gesture theory is often considered as the starting point of modern gesture research, and many important theorists in the field have built upon his work. He divided a gesture into three separate execution phases – preparation, stroke, and retraction – and introduced a taxonomy that focused on the *referent* of the gesture (Efron, 1941). It discriminated *ideographs* (gestures referring to a speaker’s thought), *physiographs* (gestures referring to the content of speech), and *pantomimics* (gestures referring to objects). He further distinguished *batons* (gestures used to stress and chunk speech parts), and *symbolic/emblematic* gestures, which lack any morphological relationship with the thing they represent. For example, the forming of a ring with thumb and index finger, signaling “OK”, can be regarded as a symbolic gesture or

emblem, since the configuration of fingers is no figural representation of any object, but merely a symbol by convention.

The second shift in gesture research started with Kendon in 1972, who focused on gestures “not as embellishments or elaborations, but as integral parts of the processes of language and its use” (McNeill, 2000, p. 9). Kendon defined a gesture as the “range of visible bodily actions that are, more or less, generally regarded as part of a person’s willing expression” (Kendon, 2000, p. 49). Kendon uses a similar classification system as Efron, but is acknowledged by McNeill (“Kendon's Continuum”, McNeill, 1992) to reference these categories with respect to their *relation to speech* and their *conventionality*. For example, gesticulations (corresponding to Efron’s batons), on one end of the spectrum, are by necessity accompanied by speech and highly unconventional and synthetic; emblems (corresponding to Efron’s symbolic/emblematic gestures) can be shown simultaneously to speech but can be understood without speech either, they have certain execution conventions but still some degrees of freedom. Sign language, on the other side of the spectrum, is designed for performance in the absence of speech, and is only understandable because it is so highly conventionalized. McNeill (2000) added two more dimensions to the classification continuum: *linguistic properties* and *semiotic character* of the gesture. He also focused on speech accompanying gestures and recognized – similar to Efron – five categories: beats, deictic gestures, iconic gestures, metaphoric gestures, and cohesive gestures.

One important difference from previous categorizations is the inclusion of *pointing movements* as a separate gesture category (*deictic*). This category is in accordance with the seminal work by Ekman and Friesen (1969), who understood gestures within the larger context of nonverbal behavior (besides facial expression, posture etc.). They labeled speech-related hand movements *illustrators*, and subdivided them into batons, ideographs, deictic, spatial movements, kinetographs and pictographs. Besides illustrators, their framework also encompassed *manipulators* and *emblems*. Emblems, in accordance with McNeill, are symbolic acts. Manipulators are hand movements which occur against the body (self manipulators, e.g. scratches, squeezes etc.), or against an object (object manipulators, e.g. manipulating a piece of clothing, a pen etc.). According to Friesen, Ekman & Wallbott (1979, p. 99) “these three classes include all hand movement except those times when the hand moves simply to establish a new position of rest”.

Further gesture typologies exist (e.g. Krauss, Chen & Gottesman (2000) who distinguish four categories: symbolic gestures, deictic gestures, motor gestures and lexical gestures), but most of them are just variants of Efron’s initial suggestion. A comparison of some major classification schemes, revealing the correspondances of categories between authors, was done by Wexelblat (1994). More recent definitions, emphasizing the purposefulness of gestures, can be found for example in Armstrong,

Stokoe and Wilcox (1995, p. 43) or Yoshioka (2005, p. 21). Armstrong and colleagues define a gesture as “an equivalence class of coordinated movements that achieve some ends”, and Yoshioka, in a similar fashion, describes gestures as “movements of hands and arms in order to achieve some communicative intent”.

As can be seen, the concept of *gesture* in psychology and linguistics is not easily defined. Rather, there is an abundance of definitions, taxonomies, and dimensions along which gestures can be differentiated. However, some characteristics of gestures seem to bridge the different approaches:

- Gestures are an act of communication, and more specifically, nonverbal communication
- Gestures are essentially movements of body parts; to some researchers they refer exclusively to arm, hand and finger movements
- Gestures are intentional (maybe with the exception of beats/batons)
- Gestures can complement or substitute speech
- Gestures vary regarding their conventionality
- Gestures can be segmented in different execution phases
- Gestures are best understood in terms of different categories. Most taxonomies have about 5 categories
- Many taxonomies have similar classes, and follow in their core Efron’s classification.

3.1.2 Gestures in Human-Computer Communication

It seems tempting to assume that the gestural communication we perform so naturally and effortlessly can be brought, by the help of technological innovation, to the domain of human-machine communication. In the best case, gesture interaction could be a high bandwidth communication channel which is intuitive to use, in the sense that we can tap into previously gained skills and knowledge, and do not have to learn a new interaction language. The definitions of *gesture* found in the HCI literature are indeed rather similar to the ones found in linguistic and psychology literature. For example, Kurtenbach and Hulteen (1990, p. 310) formulated: “A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed.” Hummels and Stappers (1998, p. 592) employ a similar definition, stating that “a gesture is a movement of one’s body that conveys meaning to oneself or to a partner in communication. That partner can be a human or a computer.”

These definitions agree with previous ones in that gestures are seen as a communicative act, and that they are performed through bodily movements. The new aspect is

the *information content* or meaning of the communicative act, while previous definitions only included an intent or willingness to communicate. Interestingly, in Kurtenbach and Hulteen's definition, a keystroke is given as a counterexample of a gesture, since the movement per se does not carry any meaningful information. The question arises, if a simple tap on a screen, e.g. on a virtual keyboard, can then be regarded as a gesture. The majority of gesture libraries count a *tap* (= a brief touch) as the simplest gesture to be performed on a 2D interactive surface (see below). Dan Saffer, in his 2008 book "Designing Gestural Interfaces" explicitly includes a touch in his definition of a gesture in the HCI context: "A gesture, for the purpose of this book, is any physical movement that a digital system can sense and respond to without the aid of a traditional pointing device such as a mouse or stylus. A wave, a head nod, a *touch*, a toe tap, and even a raised eyebrow can be a gesture" (Saffer, 2008, p. 2). Saffer's definition, contrasting previous ones, does not imply a communicative intent. For him, the act of walking through a photoelectric barrier (e.g. to trigger flushing in a public toilet) would count as a gesture, while using a stylus to enter symbols on a mobile touchscreen device would not. Again, it seems to be difficult to arrive at a generalized and standardized definition of gesture in the HCI domain, since these definitions are hard to reconcile.

3.1.3 Own Definition of Gesture

We therefore confine ourselves to defining what a gesture refers to within the context of this present research. Contrary to Saffer's view, we hold a gesture to be an intrinsically communicative act, also vis-à-vis a technical system. A human's movement, which is only accidentally captured by some sensor technology, without any intended meaning, such as for example the capturing of body movements by a surveillance camera, eventually triggering an alarm, is not a gesture. On this level, a gesture in HCI systems corresponds to a gesture in human-human communication. However, we acknowledge, along with Hummels, Smets and Overbeeke (1998), that human-computer communication is also fundamentally different from human-human communication. Sometimes, using a computer better compares with tool use than with human conversation. Or, as Walker put it: "When you're interacting with the computer, you're not conversing with another person. You are exploring another world" (J. Walker, 1990, p. 443). Yet, the user still intends to convey meaningful information, even if the information is as simple as a tap.

We thus define a gesture in the context of HCI (rather close to the definitions of Kurtenbach & Hulteen and Hummels & Stappers) as: "*a coordinated and intended movement of body parts to achieve communication. The information which it contains is specified by the configuration of body parts, the speed and direction of the movement and must be interpretable by its receiver*" (e.g. Stössel, et al., 2010, p. 271). This definition

stresses the communicative act and the intended transmission of meaningful information. It comprises static and dynamic gestures. Furthermore, a receiver capable of interpreting the movement patterns is presupposed. In contrast to Saffer's definition, no input device restrictions are given: a gesture can be sensed remotely through cameras, but also through devices such as a Wii control, a touchscreen, or stylus. Including touchscreen interaction but excluding a stylus seems not convincing.

The question of drawing boundaries to gesture interaction, however, still exists. Is all input to computers not a meaningful movement and thus in principle a gesture, even if mediated through input devices such as mouse or trackball? The important distinction we introduce is that *the movement itself needs to contain the information*. The finger movement involved in simple key presses or mouse button clicks does itself not carry any relevant information, and thus does not account for a gesture. The movement of the mouse itself, however, e.g. to reposition the cursor to a new on screen object, is regarded as a gesture (a *deictic* gesture). We argue that in case of touches on a sensitive surface, the direction of movement, and thus the positioning of the finger, carries meaningful information with regard to the *interface objects* (buttons, icons, etc.), and can thus be regarded as an *indexical* gestural act, or *deictic gesture*. The finger movement and positioning in case of key presses and mouse clicks is only meaningful with regard to the *interface device* and the real information is communicated to the computer only indirectly, separated from the original movement.

The definition given above delineates the understanding of a gestural interface for the scope of this research without claiming universal validity. Since it explicitly leaves space for various gesture-controlled systems, a categorization scheme for HCI gestures, along with suitable classification dimensions will further clarify what exactly is meant by gesture technology.

3.2 Classification of Gestures for HCI

In order to differentially analyze which gestures are better suited for older users or preferred by them (see Chapter 6), we are in need of a taxonomy which breaks up the gesture space into meaningful categories. However, since existing applications and their enabling technologies are so versatile, classification systems are not easy to establish. Gesture classification systems for HCI are also different from the previously presented gesture taxonomies for human communication (Section 3.1.1), since technological constraints need to be considered as well. In the following, existing classification approaches are reviewed and a new integrative approach is presented.

The proposed classification seeks to reconcile existing approaches and provides the basis for gesture analysis in the empirical studies.

Figure 3-1. Overview of gesture classification systems in HCI, chronologically ordered.

A) Taxonomy adopted from Quek, 1995; B) Taxonomy adopted from Pavlovic, Sharma & Huang, 1997; C) Taxonomy sketched after Hummels, Smets & Overbeeke, 1998; D) Taxonomy sketched after Quek et al., 2002; E) Taxonomy adopted and abridged from Karam & Schraefel, 2005; F) Taxonomy sketched after Saffer, 2008; G) Own three-dimensional classification system for gesture control in HCI

3.2.1 Existing Classification Approaches

Saffer (2008) distinguishes between two basic types of gesture interfaces: *touchscreen* or *free-form* (cf. **Figure 3-1 F**): While touchscreen interfaces require the user to be touching the device directly, free-form gestures use controllers, gloves, or even only the body, sensed by cameras and other tracking systems, as input. As there are fundamental differences between the gestures that can be performed on a 2-dimensional surface and gestures performed in 3D space, the labeling Saffer uses is not convincing: touch-surfaces are not necessarily screens, 3D gestures can also be very constrained - in contrast to the degrees of freedom implied by *free-form* -, and the border between contact with a device (may it be a touchscreen or a controller) and contact-free gesturing is blurred. We prefer to refer to the more clear-cut categories of *2D gestural interfaces* (all gestures occurring in a plane on a sensing surface, or otherwise tracked) and *3D gestural interfaces* (all gestures performed in three dimensions, irrespective of sensing technology).

Hummels, Smets and Overbeeke (1998) distinguish three classes of gestural interfaces (cf. **Figure 3-1 C**). The first class consists of pre-defined *symbolic commands* (also called emblems). This type of interface is still prevalent in many gesture interaction systems, but grounded more on technical constraints rather than human capabilities. The *gesticulations*, e.g. gesturing usually accompanied by speech, are much more natural and potentially intuitive to use. Yet, this type of gestures is hardly found in human-computer interfaces, because of several caveats. First, the three-dimensional nature of these gestures cannot readily be applied to two-dimensional touch interfaces. Second, the recognition of particular gestures from a continuous stream of bodily movements is still a major technological challenge, and no mature systems exist yet. Third, as there is no finite set of commands that can be learned, it is very unclear how the user should get an understanding of what the technological system is able to recognize, and what interaction options are at disposal at a given moment. The third class of gestures in their taxonomy is *act gestures*. These are gestures dedicated to a particular purpose, e.g. *descriptive* gestures (form descriptions of objects), *manipulative* gestures (transform objects), *utilization* gestures (use objects) and *indicative* gestures (select objects) (Hummels, et al., 1998). This subdivision reflects certain categories already found in human gesture classification systems, such as deictic gestures (=indicative), illustrators, pantomimic and iconic gesture (=descriptive) or manipulators (=manipulative gestures). However, to what extent these act gestures differ from the first class of *symbolic gestures*, is not totally clear. They seem to be less constrained and conventionalized than emblems, but more defined than gesticulations.

Hummel and colleagues take reference from the classification system proposed by Quek (1995), who first introduced act gestures as opposed to symbolic gestures as the two basic categories in gesture classification (cf. **Figure 3-1 A**). Symbolic gestures

have an arbitrary relationship between movement and interpretation ("a kind of motion shorthand", Quek, 1995, p. 519), whereas in act gestures, the movements relate directly to their intended interpretation. To give an example, a smoker performing the motion of lighting a cigarette in his mouth, indicating he is in need of a lighter, would be a *mimetic act* gesture. Mimetic acts, characterized by their iconicity, are to be distinguished from deictic act gestures (Quek, 1995).

Missing in Quek's system are the manipulative gestures, which were introduced in an adaptation of this taxonomy by Pavlovic, Sharma and Huang (1997). The major distinction in their taxonomy is between gestures that have an inherent communicational purpose (*communicative gestures*) and gestures that directly act on objects (*manipulative gestures*, **Figure 3-1 B**). Communicative gestures can then further be divided using Quek's classification system. This sharp boundary between communicative and manipulative gestures blurs, however, if we think for example of a point-to-select, or tap-to-select gesture. Pointing towards an object in order to select is a deictic gesture, and as such has a communicative intent. However, tapping an object on a touchscreen in order to select it (which would be the equivalent to pointing for surface gestures) could also be seen as a basic form of object manipulation, similar to doubletap-to-activate. Quek and colleagues (2002) later adopt this basic distinction between manipulative and communicative gestures, which they refer to as *semaphoric gestures* (**Figure 3-1 D**).

Yet another rather technical oriented classification system is proposed by Karam and Schraefel (2005). Besides gesture styles (which can be for example deictic, gesticulation, manipulation, semaphores or sign language gestures), the authors introduce classification categories along application domains, enabling technologies and system responses (**Figure 3-1 E**). While the differentiation according to the used technology has a large impact on the types of gestures that can reasonably be produced – one just has to compare the finger gestures enabled on a multi-touch smartphone with the hand gesture performed with a data glove or Wii controller – their distinction between *perceptual* (= non-tangible, remote sensing) and *non-perceptual* technologies (involving tangible sensors) is not fine grained enough. For example, according to Karam and Schraefel's taxonomy, a mouse, touch screen or data glove would all fall into the same "non-perceptual" category. Furthermore, as application domains are potentially endless, a classification of gesture-systems along this dimension seems to be of limited use.

3.2.2 A New Classification Approach

Comparing these different taxonomies with the ones derived previously for human gesture, some differences can be observed. First, the category of gesticulations (or,

beats/batons), which occur only together with speech does not make sense in the HCI context.

Second, since interaction with technology is not always understood best as a communicative act, but sometimes is more analogous to tool use, the introduction of a new set of *manipulative* gestures becomes necessary.

Third, the category of *deictic* gestures, which was not part in all linguistic/psychological taxonomies, is essential for gesture-based HCI systems. This type of gestures has been prevalent since the earliest gesture-based interface prototypes (e.g. Bolt, 1980).

Fourth, the space in which the gestures occur gains importance. While human gesturing is naturally three-dimensional and only restricted by biophysical limitations, the sensor technology needed in HCI brings along some constraints. Saffer, as well as Karam and Schraefel, have expressed this in their classifications, but instead of following their proposed dichotomies, we consider the gesture space as a continuum of varying degrees of freedom (**Figure 3-1 G**). Least constrained (i.e. with most degrees of freedom), are gesture systems that recognize three dimensional hand gestures remotely. Slightly more constrained are systems that still allow for three-dimensional movements, but introduce movement constraints though sensor technology mounted to the hand, e.g. data gloves, or a Wii-controller. 2D interfaces have still fewer degrees of freedom and more constraints of movement.

Fifth, as was found already in previous human gesturing classifications, a distinction between gestures according to the *referent*, or intended meaning, is of value. However, granularity of categories, the exact allocation of specific gesture types to the categories, and their labeling differ in the literature (**Figure 3-1 A-F**). Besides the already mentioned *manipulative* gestures and deictic (also labeled *indicative* or *indexical*) gestures, there seem to be at least two more categories that can be distinguished reasonably well: *iconic gestures*, which outline objects, depict forms, re-enact functionalities, and thus have a directly visible connection to the objects they refer to. In contrast, *symbolic gestures*, which have an arbitrary link between the motion pattern and its referent, can only be interpreted through the knowledge of conventions. For example, the form of a question mark (?) is associated by culture and convention to “not knowing”, and beyond, also to “need for help”.

This distinction of indexical, iconic and symbolic gestures (besides the manipulative gestures) goes back to the semiotic distinction of signs that was introduced by Peirce (1982). For the field of HCI gestures, we can place these types of gestures along with manipulative gestures on a continuous dimension of object relation concreteness (**Figure 3-1- G**).

In contrast to the alternative gesture classification systems in HCI, this system does not postulate selective and mutually exclusive categories, but rather understands the categories as exemplars along a spectrum which partly overlap, and which leave space for intermediate or mixed classes. Some gestures might be at the border between two classes, and could refer to either of them or both. As discussed before, a “tap-to-select” for example has some indexical component, but could also be seen as a very basic object manipulation.

The new classification approach includes yet a third dimension, which has been disregarded in previous systems, but is very important not only from a recognition point of view, but also from a user-centered standpoint: the complexity of the movement patterns. A gesture can be as simple as a tap on a virtual object, but it can also involve several strokes, several fingers, or hands. Again, gesture complexity can vary along a spectrum, ranging from a simple gesture with one finger involving only one stroke, to gestures that combine both hands in an elaborate movement pattern. These gestures can exist for all gesture space levels (2D and 3D) as well as for all gesture classes (e.g. a symbol or a direct manipulation). This three-dimensional gesture taxonomy (**Figure 3-1 G**) is intended to be generalizable to all gesture-based HCI systems, and each gesture can potentially be anchored within the grid.

3.3 Technological Aspects of Gesture Control

Even though this research has the pretension of investigating touch & gesture input for elderly users on a broad and basic level, we had to restrict the research space on two dimensions. First, only two-dimensional surface gestures were investigated, and the potential of three-dimensional gestures was not fathomed any deeper. The reasons for this decision were twofold: from a user point of view, the gesture systems involving touchscreens are by far the most abundant, they are most likely the ones that older users will get in contact with and the results are most relevant to devices currently on the market or currently being developed. From a technological point of view, these systems are also the most mature; they exist in mass consumer products and not only on prototypes or niche products, and the least technical difficulties related to gesture registration and classification are expected.

The second restriction was taken with respect to gesture complexity: only finger gestures were investigated, and these were also usually restricted to at most two fingers. This restriction naturally followed from the focus on mobile devices. Their limited screen size did not allow for true multi-finger or even multi-hand gestures. The concentration on mobile devices was grounded on similar arguments as the

restriction to 2D surface gestures: they are most readily available and of greatest relevance to the older users.

Within the three-dimensional gesture interaction space established in **Figure 3-1 G**, this research covers only a certain section, and leaves space for further investigations, such as the potential of larger touch-based gesture systems (tables, walls, ...) or even three-dimensional gesture control systems for older users. On the dimension of gesture class, however, we cover the whole spectrum from manipulative to symbolic gestures. The surface gesture systems under investigation are characterized further in the next section, in which technical aspects and limitations of these systems are discussed, and what they mean with respect to a senior-friendly interaction. A brief historic account of the development of multi-touch technology over the past decades can be found in **Appendix A.7**.

3.3.1 Framing Touch Technology

The variety of touch-sensitive systems ranges from palm-size to wall-size devices, some are operated by stylus, some by fingers, some by hand, some are horizontally mounted, some vertically or at oblique angle, some recognize merely touch points, some complex gestures. A classification scheme for these systems with focus on the implemented gestures has been presented in Section 3.2.2, but other technical aspects have not yet been covered. Since the technical implementation determines to a large degree the capabilities of the system and the way the user interacts, some of the key design differences in multi-touch technology are discussed in the following, mostly in the form of contrasts. In a similar way, these contrasts are also described by Buxton (2011).

Tablets vs. screens. Many of the early (multi)touch devices were developed just as *input* devices for a PC, as an alternative to the traditional mouse and keyboard, and realized in the form of pads or tablets, while today's touch input is realized mostly in form of *touchscreens*. The major distinction is the one of directness (see Section 2.5.2): while on a screen the user is able to manipulate the visible objects directly, on a pad or tablet the input occurs at another location from the visual output, which comes usually at cognitive translation costs. This distinction also guides the naming of the systems: whenever the touch surface is not collocated with the screen, we speak of touch pads (usually smaller) or touch tablets (usually larger). Since touchscreens abound in modern devices and feature the benefit of direct manipulation, the thesis will focus on their investigation and disregard pad or tablet input.

Discrete vs. continuous. Touch-based systems often differ on the types of input events they can handle. Until recently, the majority of touchscreen systems known to the public (e.g. those installed in ATMs or ticket machines) made only use of discrete

events, such as the press of a virtual button. Only since a few years, manufacturers have realized that the richness of information that can be conveyed by continuous input, e.g. moving the finger in a certain pattern on the screen, gives a tremendous boost to the usefulness of touchscreen technology. This distinction is not analogous to the difference between single- and multi-touch screens, even though in reality systems capable of multi-touch input are usually also capable of sensing continuous input. Since only systems which can handle continuous input allow performing gesture patterns, and consequently only these systems bear the potential of leveraging these gestures for a less cumbersome and more intuitive use, these are the only ones investigated in this thesis.

Palm vs. wall. Multi-touch devices have been designed in all kinds of sizes. The latest generation (6th, released 2010) of the Apple iPod Nano includes a multi-touch screen of less than 4 cm (1.54 inches), while Perceptive Pixel ships multi-touch screens more than 100 times larger (e.g. 160 inches). Naturally, the size of the interface determines also the type of interaction. Even though the technology is very scalable and solutions exist for various intermediate sizes, it might help to think of these systems in five categories:

- *Sub-palm sized:* Interaction occurs mostly only with one finger, and a reduced set of gestures.
- *Palm sized:* Gesture possibilities are more diverse and multi-finger input is possible, but usually is restricted to two, at most three fingers simultaneously.
- *Tablet sized:* Usually features a rich set of finger gesture input, and has few constraints on simultaneous finger input. Can also incorporate gestures for two-hand interaction.
- *Table sized.* Stationary, and usually horizontally mounted. Gestures that involve fingers from two different hands are common, but also gestures that involve the palm instead of singular fingers. The large size of the display allows that several users perform gestures simultaneously, opening up a new spectrum of collaborative gestures and tasks. Furthermore, the horizontal and fixed mounted nature of these devices creates the possibility to interact with other physical objects which can be placed on the surface and sensed by the system.
- *Wall sized.* Still larger multi-touch systems are usually vertically mounted, and size restrictions are only given by the limited reach of the human arm. These systems are especially designed for collaborative work and complex information visualization. These systems are yet relatively rare and only used in very specific contexts. If the screens grow so big that the information displayed cannot fully be captured any more by the user's visual field at arm length distance, the multi-touch interaction becomes ergonomically problem-

atic. Some systems of such size have therefore tried to capture also touchless gestures in addition to the touch interaction (or switched completely to free form gestures), so that the user can interact with the system still at a more comfortable distance.

The research in this thesis concentrates on the multi-touch technology which is most readily available and has the broadest application base. Therefore, only touch- and gesture systems which are palm sized or tablet sized are under investigation.

Horizontal vs. vertical. The angle of mounting a multi-touch device also sets boundary conditions for the interaction style. Fatigue of the arm for example is much more of concern for gestures performed on a vertically mounted touchscreen than for horizontal systems, where wrists or elbows can be supported by the surface. On the other hand, a horizontal viewing plane is usually not supportive of an ergonomically convenient neck and body posture. Most mobile touchscreen devices are therefore used in more natural oblique angle, in a plane orthogonal to the line of sight.

Position vs. pressure. The vast majority of current single- and multi-touch systems exploits only *whether* a contact has been established, and *where* it was established. This purely spatial information, together with derivatives of speed and acceleration, is sufficient to distinguish a huge set of gesture patterns. Some prototype systems (e.g. the early system by Herot & Weinzapfel, 1978) can make use of still richer information. The *degree* of touch can be measured through pressure sensitive systems or by sensing the area of contact (which increases with increasing pressure). Some systems have even demonstrated the ability to sense the angle of the finger relative to the surface, and by this creating a vector which can allow for sophisticated 3D manipulation of objects. The present research however only fathoms the possibilities of gesture interaction on the basis of pure contact and position information, since its application space is already vast and this information can be extracted from all commercially available systems.

Single-touch vs. Multi-touch. As the name suggests, a single touch screen can only recognize one point of contact at a time, while a multi-touch screen can recognize more than one (the exact number varies from system to system). Gesture interaction, however, is possible on both systems. Buxton makes a strong claim in favor of multi-touch by stating “if you can only manipulate one point, [...] you are restricted to the gestural vocabulary of a fruit fly. We were given multiple limbs for a reason. It is nice to be able to take advantage of them” (Buxton, 2011). In our view, it is not so much the number of contact points, but rather the question if only discrete events can be registered or whether continuous movements are possible that largely determines the richness of interaction. Even single touch

systems that can reliably interpret movement patterns performed with only one finger (e.g. a flick gesture on a picture to proceed within a slideshow) provide a new dimension of directness and naturalness as compared with static (virtual) button presses. Within the scope of this work, both single touch and multi-touch systems are being investigated.

Multi-hand vs. multi-finger. From a technological point of view, many multi-touch systems are capable of sensing an unlimited number of contact points. Most systems cannot, however, distinguish whether the contact points are established by multiple fingers of the same hand, by multiple fingers from different hands, or even by fingers from multiple users. Especially on larger screens designed for collaborative work, this becomes problematic, if for example a second registered touch might be either from the same user, and thus maybe denoting the edges of a square selection area, or from a second users, and thus maybe denoting the positioning of an independent second cursor (an exception is the Diamond Touch system developed at the Mitsubishi Electrical Research Laboratories, 2011). This problem is not so much of relevance for the present research since we only investigate small single-user systems. However, when it comes to designing suitable gestures, the matter of single hand vs. two hand gesture execution must be considered.

Finger vs. Stylus. While most of today's multi-touch systems are operated by fingers, some are also designed to be operated with a stylus or work with both finger and stylus. Obviously, the stylus has a much smaller touch area, and therefore allows implementing smaller touch targets on the screen when space is limited. On the downside, the interaction becomes less seamless, the stylus can get lost, touch targets remain visually small, and multi-touch cannot be realized by using a stylus. For all these reasons, a stylus-based interaction does not seem appropriate for older users, and therefore this work is concentrating on technology which can be operated just by finger movements on the touch-sensitive surface.

Multi-touch technology can occur in a multitude of facets, and maybe it is more appropriate to think of it as an umbrella term for different input paradigms rather than a homogenously defined input style. Among this plethora of technical systems, this research is concerned with only a specific type of system, which, however, is believed to be the most readily available and most appropriate for this context. The touch and gesture input systems that were investigated were only *touchscreens* capable of sensing *continuous position data*, operated directly with one or more *fingers simultaneously*, and mostly of *handheld* size.

3.4 Chapter Summary

At the beginning of this chapter, the foundations for understanding gesture interaction in HCI were laid by characterizing what qualifies a movement as a gesture in human nonverbal communication. Several definitions and taxonomies of gesture in human communication were presented and compared, and key common features which bridge the different approaches were extracted. The gist of human gesture was subsequently compared with the definitions of gesture in human computer interaction. One important difference is the stress on the information contained within the movement, which needs to be extractable by a technical system. An integrative definition of gesture in HCI was proposed, and its scope was discussed. Subsequently, a series of classification systems for gesture interaction in HCI were compared, and their respective differences and shortcomings were pointed out. On these grounds, a more general classification system was proposed, which assumes continuous dimensions rather than discrete states, and spans the space of gesture interaction systems along the dimensions of gesture complexity, gesture space and gesture class. The focus of the research was put on touchscreen technology and two dimensional surface finger gestures. It was argued that because of availability, flexibility, relevance and technological maturity, these systems currently offer the greatest potential of improving technology interaction for older users. A series of technical aspects were discussed and evaluated regarding the implications for a senior friendly design. On this basis, the focus on a certain type of touchscreens, of handheld size, capable of sensing continuous input, and operated directly by fingers, was justified.

4 Gestural Interaction for Elderly Users

The second chapter portrayed the older user, pointed out the characteristic changes in sensation, perception and cognition that accompany the normal aging process, and related these to the problems older users have in everyday life when dealing with interactive technology. The third chapter then changed the perspective towards the technological side of interaction. It provided a framework for understanding and categorizing the emerging touch and gesture-based interaction styles. We are now confronted on one side with an interface evolution, which has become the dominating paradigm on the smartphone market (and increasingly so also for all mobile devices), and on the other side, with the steadily progressing demographic shift, which puts older adults more and more in the center of technology usage scenarios. In Section 2.5 the multitude of problems that older users are still facing with regard to technology usage, and in particular with regard to mobile devices, has been pointed out. Given the range of usability problems with existing technology, and given a new paradigm which is connoted to be more “natural” and “intuitive” to use, the question which is at the heart of this research emerges automatically: does this new technology bear the potential to ease the interaction problems particular for the older users? That is, can it contribute to reduce some of the usability issues this older generation is currently facing when interacting with technology?

To answer this question, this chapter starts with a review of existing solutions and research concerning older users and gesture-based interaction. Based on these insights, some theoretically derived arguments in favor of and against the use of such technology for older adults are presented, which in turn are used to derive the research questions, hypotheses and research agenda of this research.

4.1 Existing Solutions

One area where gesture technology has successfully improved the lives of older adults is geriatrics. Various reports exist of Nintendo’s Wii gaming console employed in senior and retirement centers, day care, or rehabilitation centers as means to bring a bit of physical activity back to people who have been forced into more sedentary lifestyles. Requiring few financial and organizational resources, therapists can now combine physical activities with play and competitive motivation as well as social interaction. The system allows seniors to compete e.g. in bowling, boxing,

tennis and other type of games by performing corresponding gestures with the Wii remote control. The benefits of such activity on muscle strength and balance have been scientifically investigated and documented (e.g. Homma, 2009; Nitz, Kuys, Isles, & Fu, 2010). While it is easy to see that the playful movements involved in these types of games are better for older people's physical wellbeing compared with sitting in an armchair, it is much harder to see whether this technology can actually help to overcome any real interaction problems with everyday technology required to perform essential activities. Since these are two totally different fields, all that can be transferred from these studies is the fact that older adults, although often skeptical at first, can embrace and enjoy such new technology if they perceive a benefit, and if they can draw satisfaction from mastering the tasks.

Studies where touchscreens have been employed as information and communication systems dedicated specifically to older and disabled users (e.g. Holzinger, 2002; Holzinger, 2003; Otjacques, Krier, Feltz, Ferring, & Hoffmann, 2009) generally report a good acceptance and user friendliness of such systems. However, the systems under investigation in these studies were tailored especially to older people, and also have an adapted, largely restricted graphical user interface. Therefore, the impact of the input modality per se on the whole user experience is hard to determine. Furthermore, the system evaluation in these studies is only based on subjective reporting and small user samples (N=5-12). Comparative studies which compare touch-based interface devices and other input devices on more objective grounds, such as time on task or accuracy (e.g. Armbruster, et al., 2007; Mahmud & Kurniawan, 2005; McLaughlin, et al., 2003; McLaughlin, et al., 2009; Murata & Iwase, 2005; Murata & Takahashi, 2008; Rogers, et al., 2005), are not univocal (see Section 2.5.2). Still, the general benefit from the direct nature of touch interaction, namely requiring less sensorimotor transformation can be observed in these studies. However, they all test only the single touch aspects and usually involve simple homing tasks for virtual buttons. The studies disregard the potential of multi-touch and gestural interaction to leverage this technology for building more easy-to-use systems for the elderly. In fact, to our knowledge, there is almost no literature which has investigated the use of gesture interaction for older users.

An exception is the "SharePic": a multi-touch, multi user, gestural tabletop system developed by Apted, Kay and Quigley (2006) especially for elderly users to edit and share digital photographs. Their design process was guided by design principles tailored to the interface and user group, e.g. to ensure easy scalability of interface objects (and thus accommodating for visual impairment), to focus on learnability and memorability, to consequently minimize interface elements, and to strive for predictability and consistency. Besides tapping and dragging pictures, the interface includes specialized two-hand gestures for rotating, resizing and copying of pictures. The

system was evaluated through qualitative video analysis of six pairs of older users (between 63 and 81) working on two tasks, involving sorting pictures and creating a digital postcard. Results showed that the older users had generally no problems to learn the gestures for generic actions like moving, resizing and rotating a picture, but more specific actions were difficult to learn. Especially the copy gesture proved to be difficult. This gesture, however, had been implemented similarly to the resizing gesture of other gesture-based systems (placing two fingers on an object, then moving one finger away while the other stays stationary). For other actions such as *erase*, particular interface objects such as the “black hole” needed to be interacted with, which makes it difficult to generalize the findings of this study to other gesture based systems. In addition, the only task tested, arranging digital pictures (the only task the system was designed for), is not usually considered to be among the most frequent tasks that older people would want to perform with a digital device. A more general application domain, a more structured gesture design process, and better performance measures are necessary in order to make any valid conclusions if this interaction paradigm is beneficial to older users on a general level.

In a different study, a comparison of a traditional mouse interface with a pen-based interface was conducted with thirty older users (between 60 and 83) by Arrias-Torres (2006). The pen-based interaction style involved straight line gestures to be executed on top of screen objects in different directions, in order to issue commands such as open, close, copy, paste, undo or delete. Even though the results of this study are only given by subjective measures (the SUS, system usability scale, Brooke, 1996) and only descriptively, it is remarkable that novice older users rated the pen-based interface, despite its rather arbitrary gesture patterns, better than the established mouse interface. This result is in accordance with earlier research that has documented the benefit of direct interaction with a pen over a mouse (Charness, et al., 1995; Charness, et al., 2004).

Leonardi and colleagues developed a gestural touch-based device called the “MobiTable” to be installed as digital assistant in the homes of elderly people. (Leonardi, Albertini, Pianesi, & Zancanaro, 2010; Leonardi, Mennecozzi, Not, Pianesi, & Zancanaro, 2008) It was designed around the concept of familiarity, pursuing a participatory design process that included older people in contextual enquiries, focus groups and workshops. Familiarity was used as the driving concept behind the design of the gestural interaction patterns: “The physical actions needed to perform a task should recall motor patterns used in the real world, for instance scrubbing with the finger an object to erase or consume it” (Leonardi, et al., 2008, p. 4). The main function of the MobiTable was to support older adults in social networking, and a specific GUI for sharing information and communicating with peers was installed. A qualitative evaluation study with 15 older people aged 62 to 93 revealed a general

positive adoption of the gestural interface, with participants quickly mastering the basic mechanics of the gestures. Interestingly, the authors report that participants changed from an initial strong preference for the use of the stylus to a slight preference for the finger as they became more comfortable with the device. Furthermore, the authors stress the value of iconic gestures as very engaging and motivating especially for users with high computer anxiety (Leonardi, et al., 2010). From their observations, Leonardi and colleagues conclude that special care should be taken 1) to allow fine-tuning of timing parameters of gesture interaction (e.g. the time between touch and release), which, if possible, should be adaptive to the user because of the huge variability, 2) to avoid taps which are not object-centered, and 3) to avoid overloading a target object with multiple gestures (e.g. touch and drag), since users might get easily confused. While these results are encouraging for a positive impact of touch and gesture control on the user experience of older people with technology in general, a more structural approach to gesture design as well as a more rigid and quantitative evaluation is still needed.

Taken together, these studies provide a first indication for a successful employment of touch-and gesture controlled technology to the benefit of older adults. However, these studies are only conducted on specially designed prototypes, and lack a solid evaluation, including larger samples and quantitative behavioral usability metrics. Furthermore, the tested applications such as picture sorting and social networking were rather specific. Research has yet to show if gesture technology can also be used by older adults for generic tasks and for everyday devices, such as mobile phones, remote controls, personal computers etc. In particular, the use of gesture control also on small screen handheld devices, where it is currently most readily available, still needs to be scientifically investigated.

4.2 Potential Benefits and Drawbacks of Gesture Interaction

From a theoretical point of view, touch- and gesture based interaction on mobile devices presents a number of advantages and disadvantages for older users. Some of these are related directly to the nature of touchscreen interaction, others refer to the potential of performing gesture patterns. The following section compares the potential benefits and potential drawbacks of such technology, which has been delineated previously in a similar way in Stöfel & Blessing (2009).

4.2.1 Potential Benefits

- **Indirect size advantage.** With the introduction of highly responsive touchscreen technology, it was possible to replace the previously dominating method of operating a mobile phone – a grid of small hardware keys – with a large area touchscreen, that allowed for single- or multi-touch interaction. Consequently, assuming equally sized devices, more screen space is available, and all visual information can be displayed bigger. This simultaneously addresses two of the most frequently named interaction problems of older users: buttons too small to press, and text too small to read (see Section 2.5.1).
- **Unique labeling.** The size advantage also provides the possibility of unique labeling. Because of size restrictions, labels of physical buttons on mobile devices are not only small and thus hard to decipher, but it is common practice to give multiple labels to the same limited set of physical buttons, which have different functionalities in different modes. Typically, buttons on mobile phones are labeled with a number (for dialing) and a series of letters (for text writing), but other symbols for diverse functions are included as well (e.g. screen locking, direct access to functions etc.). Not only are these hard to read, but also confusing to older users. Virtual buttons on an all-screen device do ideally only appear when relevant for the current application, and only with the currently relevant label.
- **Adaptable GUI.** As virtual buttons replace hardware keys on touch-enabled devices, they are – in principle – freed from hard constraints on size and form. It is theoretically possible to load different profiles on a device, which could feature differently sized, labeled and positioned interface objects on the screen, reflecting the needs and capabilities of the target user. For example, the same smartphone hardware, whose homescreen features a grid of sixteen rather small icons to access a large variety of functions, could also show nine applications icons for intermediate users, or four icons for beginners with direct access to only the most basic phone function, which might be appropriate for many older users. Currently available devices do not yet exploit this potential.
- **Direct interaction.** The direct nature of interaction of touch-and gesture control has already been discussed in Section 2.5.2, as well as the general benefit of direct interaction styles for older users (e.g. Murata & Iwase, 2005).

- **High bandwidth interface.** A simple touch on an icon on the screen does not convey a wealth of information to the system. The only variation possible is the duration and frequency of touch (e.g. a tap, double tap, or long press). This results in a limited interaction spectrum. Using gestures, the information bandwidth is much higher. With a single gesture, a map could for example be rotated by a certain angle and magnified by a certain degree at the same time. Compared to a key-based interface, “a single gesture can be equivalent to many keystrokes” (Rhyne, 1987, p. 138).
- **Intuitive gestures.** The mapping of a gesture pattern to the function it triggers might be arbitrarily chosen, and as such provide little benefit to an older user. It would take effort to learn these mappings, comparable to learning to control a device over menus and buttons. If, however, familiar symbols or movement patterns are used as gestures, the older users could tap into deeply rooted cultural and sensorimotor knowledge resources, and thus facilitate effortless interaction (Stössel & Blessing, 2009). Less domain knowledge or product expertise would be required, since the users can unconsciously apply knowledge previously acquired, which is a precondition for intuitive use (Blackler & Hurtienne, 2007; Mohs, et al., 2006; Naumann, et al., 2007). For example, while the process of deleting or confirming an item on a traditionally operated cell phone involves opening a hierarchical menu with the help of soft keys and selecting the appropriate command through the navigation keys, a gesture-based approach could simply involve performing a “crossing out” or “ticking off” gesture on the item in question. The traditional approach will only be intuitive to people who have some computer literacy and experience with similar systems, while the gesture-based approach will be intuitive even to people with no prior domain knowledge.
- **Procedural knowledge.** Not only might relevant mappings between gesture and function be easily remembered (since the associations have been learned before in different contexts), their execution might also be performed with relative ease. It has been documented that procedural knowledge, such as the correct execution of a hand motion pattern, is relatively robust to age-related decline compared with other memory systems (e.g. Mitchell, 1989). Furthermore, it was shown that older experts in a certain skill show reduced, if any, age-related declines in the speed or the efficiency with which they perform skill related tasks, while normal age-related declines were found on general tasks and unfamiliar materials (Krampe, 2002). If the gesture patterns are selected in a way that older adults are already highly trained in their execution (e.g. “ticking off” an item, “crossing out” an item, drawing an alphanumeric

symbol), older users should not suffer substantial difficulties recreating them on a technical device.

- **Balance of hedonic and pragmatic qualities.** As mentioned previously, devices which are purely designed with ergonomic requirements for older users in mind might run the risk of being highly functional, but at the same time stigmatizing the user and having little hedonic qualities. Older users too want products which are aesthetically pleasing, and fun to use (Pohlmeyer, et al., 2009), but without giving up good ergonomics and usability. In a similar argument, but this time from the manufacturer's instead of the user's view, Keates and Clarkson claim that "a strategic decision is therefore required regarding the balance between the marketability of the product and the level of population exclusion" (Keates & Clarkson, 2003, p. 69). Since gesture interaction is generally believed to provide an engaging and joyful user experience, and at the same time bears the potential of efficient high bandwidth interaction, this technology might actually constitute a good balance between hedonic and pragmatic interaction qualities, or between the marketability and the level of population exclusion.

4.2.2 Potential Drawbacks

Naturally, the novel form of interacting through touch and gesture is not the holy grail of interface design, but entails at the same time disadvantages and potential risks, some of which might be particularly detrimental to older users. Norman and Nielsen (2010), have indeed called gestural interfaces "a step backward in usability", and argue that some fundamental principles of interaction design are increasingly being disregarded by gesture based systems. Some of these drawbacks which are relevant with respect to older users are presented in the following.

- **Loss of cues and affordances.** Gestures are designed to work independently from on-screen items that provide cues or affordances of how to interact with the system. For example, the well known *pinch* and *spread* gesture to enlarge or shrink a picture, is not hinted at anywhere in the GUI. Formerly, there might have been a menu entry "zoom" which had to be discovered and selected by navigating through a menu with the navigation keys, or, in a pure touch interface, there might have been a "+" and "-" virtual button on the screen, affording to touch it. While a hardware button affords to be pressed (e.g. to cancel a phone call), and menu items usually provide textual or graphical cues on their functionality, nothing hints at which gesture could be employed in which context. Norman & Nielsen (2010) call this the problem of visibility and discoverability: not only are there no signifiers how to perform a certain action, but it

is not even possible to discover them through systematic exploration of the device. Particularly older users might be easily at loss in such systems, if they have forgotten a certain interaction gesture, and the system offers no signifiers or other hints on how to perform the desired action.

- **Lack of consistency and standards.** As with all new technology which is not yet matured and for which no standards exist, manufacturers are striving to stake their claims and compete in creating the future interaction standards. This leads to a multitude of different gesture realizations for the same actions, and vice versa, the same gesture for different actions across various systems. “Touching an image can enlarge it, hyperlink from it, flip it over, unlock it so it can be moved, or whatever the developer and his whim decided” (Norman & Nielsen, 2010, p. 48). This lack of consistency makes learning much harder, since knowledge cannot be transferred easily between similar devices or systems, and this is aggravated for people that show shallower learning curves, such as the elderly.
- **Loss of feedback.** On traditional interfaces, e.g. those involving a mouse and keyboard to control a PC or a set of keys to control a mobile phone, users receive immediate haptic and acoustic feedback (“click”) on whether the finger is positioned correctly and whether a button press did occur. This feedback is lost on touchscreen interfaces, where fingers have to be positioned and moved across a uniform surface. This becomes particularly apparent in the case of gesture interaction: in case a finger movement pattern is executed on the screen and no immediate reaction of the device occurs, the user has no clue whether the gesture was performed at the right place, the right direction, the right speed or the required precision, whether the device detected some input but cannot resolve it, or whether the system answer is just taking time. Since older users have a special interest in safety and security during interaction, and are often scared of getting lost within a system or damaging a system through inappropriate use, this loss of feedback could be perceived as especially detrimental among this user group.
- **Complexity of motion.** The complexity of hand and finger motion required by gestural interaction can easily exceed the demands posed by key interaction. Pressing a physical button allows a certain sluggishness of the movement, the center of the key does not need to be precisely hit, the finger can be repositioned before the final press, speed of movement and pressure are largely irrelevant. In gestural interaction, the dynamics and the trajectory of the movement also carry information (e.g. the speed of a *flick* gesture can determine

how fast and far a list is scrolled), and often two (or more) fingers need to be coordinated in a simultaneous movement pattern to achieve the interaction goal. The literature of motion control of older adults (see Section 2.2.3) reports various age-related changes, e.g. a general increase in motion variability, a general slowing of movement, a loss of force regulation, reduced coordination, as well as reduced range of motion. Depending on the specific gesture set implemented in a system, the motor demands might present a particular burden to older users and lead to frustration, if they are not able to perform the gesture as required.

- **Risk of inadvertent activation.** Compared with an indirect interface like a keyboard, the direct nature of a touchscreen also brings about the danger of inadvertent activation. It is much easier to touch the screen unintentionally and accidentally trigger an undesired function. On modern highly sensitive touchscreens which react to the slightest contact and do not require any pressure, this can become a problem. If the system reacts too directly, and therefore is less error tolerant to uncontrolled finger positioning and movement, it could turn out be particularly difficult to use for the older generation.
- **Acceptance.** Elderly users are generally more reluctant to adopt unfamiliar technologies (see Section 2.4). As gesture input presents a radically new interface concept, they might be reluctant to make use of it, even if they scored better on usability metrics with gesture input.

4.3 Research Agenda

At the heart of this thesis is the investigation whether touch and gesture based interaction provides a benefit to older users over traditional input methods, or whether the drawbacks outweigh the benefits. Along with Fisk (2004), we consider older adults as those individuals aged 60 and above, and focus in this research particularly on the *younger old* (see Section 2.1) in an age-band of roughly 60-75. By narrowing down the age-band for the empirical studies, we strive to reduce the variability in skill and knowledge levels in this nonetheless extremely heterogeneous group. Moreover, we are interested in technology which is adapted to the needs and abilities of the healthy older population, taking into account all the changes in physiological or cognitive resources that occur during the normal aging process, but not focusing on certain pathologic states. Since pathologic conditions also occur with increasing frequency at higher ages, this was another argument to focus on the younger older

users. Lastly, the lifestyle of potential users was also taken into account. Since touch and gesture based technology is abundant on mobile devices which are currently pushed on the market, it made sense to investigate its use also within an age group which is still largely mobile: a generation which has just finished or is at the end of their professional life, but which is still rather active and socially connected. A generation that could profit from devices that allow to communicate with friends and family, that help in navigational tasks when losing one's way, that help to memorize things and to retrieve information, that provide a feeling of safety because they are always reachable and traceable, that might even be fun and informative and fight boredom and loneliness ... if such devices were only usable.

The problems of older users with existing technologies, and particularly with mobile devices, have been pointed out in Section 2.5. It is argued that to a large extend, these problems arise because technology is not designed for the special needs and abilities of older users. Amelioration, in our view, can only be achieved if the older user is put in the center of the design process. In the empirical studies (Chapters 5, 6), we therefore measure the users' capabilities, ask for user opinions and let the user evaluate certain aspects of the interaction. The insights from these studies then guide the design of a functional prototype, which is again evaluated by the older users (Chapter 7).

From the technological point of view, the new paradigm of touch and gesture control was the prime candidate to ease some of the usability issues. Just recently has this technology become the dominating input method for most modern mobile devices (see Section 3.3). Even though some hints exist that gesture and touch technology is well received among older adults (Section 4.1), there has not been a systematic and thorough investigation of the potentials and limitations of its use in this context. To this extent, the present work is discovering new scientific shores, but at the same time remains largely explorative, since no consolidated findings or theories exist on older adults' use of gestural input technologies. Several models try to explain technology acceptance or adoption of older people, but for the reasons specified in Section 2.4, none of the existing models seemed appropriate to be singled out and guide this research. Rather, the present work takes an applied perspective, focusing on the actual use of a certain technology by elderly people.

4.3.1 Research Questions

The basic question, whether gestural interaction provides an advantage to older users or whether the drawbacks outweigh potential benefits, can be broken down into the following more fine-grained questions:

1. Q1: What are the requirements of gesture technology for elderly users regarding **motor control** and **manual dexterity**?
2. Q2: To what extent, or under which circumstances, is gesture technology **intuitive** to understand and use for the older user?
3. Q3: How easy is it for older users to **learn** and **remember** gestures for controlling technology compared with conventional interaction paradigms?
4. Q4: Does gesture technology provide a **benefit** for the older user? Is this technology **accepted** by older users?

From a user's point of view, the problem is analyzed on three levels:

1. **Physical level:** At this most basic level, it is of interest if the demands to motor control and manual dexterity hinder older users to benefit from gestural interaction. In other words, it is of interest to investigate under which boundary conditions, e.g. with respect to the size of a screen, the complexity of a gesture pattern, the posture of a handheld device etc., older users can make efficient and effective use of gesture input. (Q1)
2. **Cognitive level:** On a second level, it is of interest to investigate which cognitive effort is needed to perform successful gesture interaction. We assume that gestural interaction can be performed by older users with little cognitive effort if the input model matches their mental model, and this is achieved if the gestures can be "intuitively" applied. The main research goal on this level is to find out which gestures are intuitive to the older user, and whether these might be different from the ones that are intuitive to younger users. (Q2,Q3)
3. **Behavioral level:** On this level, the actual interaction properties are of interest. Assuming that older users interact with a gesture controlled device that matches their abilities on a physical and cognitive level, how well do they fare on classical usability metrics such as efficiency, effectiveness or satisfaction? Can a gesture-controlled device outperform traditional, key-operated devices? And even if this is the case, is the technology accepted by older users, or are there other factors which keep them from actually using it. (Q4)

From a technological point of view, the problem can also be analyzed on three levels. If we understand gesture interaction as a sign system, we can investigate such a system from a semiotic point of view along the three branches of semiotics: syntactics, semantics and pragmatics (Morris, 1938/1970).

1. **Syntactics:** Syntactics is the branch of semiotics that deals with the formal properties of signs and symbols. Concerning gestural interaction, these relate to the properties with which a gesture can be formally described: the area (or

volume) a gesture covers, the number of segments a gesture consists of, its curvature, its angle, its structural complexity, the number of fingers needed for execution, etc. These properties can be directly linked to the user's motor demands (Q1).

2. **Semantics.** Semantics is the branch of semiotics concerned with the relation between the signs and their objects of reference. In the case of gestural interfaces, this leads to the question which signs are best suited to denote certain activities that a user might want to perform on a system. There is a need to investigate which gestures are easily understood, can be intuitively applied, can be learned effortlessly and memorized without trouble. These questions directly relate to the cognitive user level (Q2, Q3).
3. **Pragmatics.** Pragmatics is the branch of semiotics interested in the ways in which signs are used and interpreted. It deals with the influence of context and the effects of signs on their users. In the context of a gesture interface, the pragmatic interaction aspects comprise the level of performance within a real interaction setting, the level of interaction errors, the interaction speed, the level of user frustration or satisfaction, as well as the level of user acceptance of such technology. This corresponds to the user behavior which can be measured through classical usability metrics. (Q4).

4.3.2 Structure of the Research

Both from a user's point of view and a semiotic point of view, the research space can be divided into three areas, which roughly correspond. These three areas have also guided the research process.

Stage 1: Conceptualization. The most logical starting point was to investigate the physical properties of gestures, and how they relate to the motor constraints that come with age. Do older users have problems in performing gesture patterns, and if so, under which conditions can these problems be minimized? The vast space of possible gesture alphabets and technical realizations should be narrowed down to those which are most apt for older users. Two studies (1A, 1B) have been carried out at this stage, which are described in Chapter 5. Study 1A investigated gesture execution performance on a stationary touchscreen, with varying gesture space and gesture complexity. Study 1B, using a similar setup, investigated gesture execution performance on a mobile device. This time, the screen size was fixed, but posture was additionally varied. The objective of this stage was to answer research question 1.

Stage 2: Conceptualization II. While in the first stage only the patterns were analyzed without any underlying semantics, the second stage was concerned with the meaning of gestures. In particular, the aim was to further narrow down the possible gesture interaction space to those gestures which are intuitively understandable to older users. We were interested in finding out whether there are any common features in the conception of intuitive gestures among elderly, and whether the gestures which older people find appropriate differ from the ones that younger users prefer, or from the ones which current devices use. Two studies have been carried out in this stage (2A, 2B), which are described in Chapter 6. The objectives of stage two were to answer research question 2, and to identify a gesture set which is empirically justified as a promising candidate for a gesture interface adequate for older users.

Stage 3: Implementation. Stage one and two can be regarded as the conceptualization stages, in which on the one hand insights on the relationship between older people and technology use were gathered, and on the other hand a promising set of gesture was established. Stage three focuses on synthesis, rather than analysis. In this stage, a functional prototype was developed based on the findings from the two previous stages. To this end, an Android smartphone was programmed to mimic a typical graphical user interface, to provide a series of basic functionalities, to respond to a custom designed set of interaction gestures, and to create a continuous logfile of user input for subsequent analysis. Chapter 7.1 describes the development process and the features of the prototype.

Stage 4: Evaluation. In stage four, the functional prototype was evaluated with 32 older users. The touch- and gesture interaction paradigm was compared to the more traditional input method of hard keys on a mobile device. By comparing these two interfaces, advantages and disadvantages of the gesture interface could be quantified with subjective and objective measures of usability. Among other criteria, the time-on-task, the number of redundant interaction steps and the frequency of screen changes were assessed as measures of efficient interaction. Both interfaces were also evaluated with regard to their perceived effectiveness, familiarity, controllability, learnability, intuitiveness, perceived error rate and helpfulness. Even though the older users' interaction with the system was of limited time, the objective of stage 4 was to find at least partial answers to question 3 and 4. The evaluation study is described in Chapter 7.

5 Gesture Syntax and Motor Demands

In the first stage of conceptualizing appropriate gesture input for older adults, it was of interest to investigate how well older adults are able to produce gesture patterns in varying conditions. Two studies were devised to this end: in study 1A, gestures had to be reproduced on a stationary touchscreen, while the size of the gesture space and the complexity of the patterns were varied systematically. In study 1B, a similar set of gestures was tested on a mobile device. This time, there was a fixed screen size, but the difficulty of the gestures was varied as was the operation posture of the device. In addition, we were also able to assess multi-finger gestures.

The basic assumption behind these studies was that an age-related decline of motor control and manual dexterity could be detrimental to the faultless execution of finger gestures. These studies were designed to investigate whether this was indeed the case, and which boundary conditions, or gesture patterns, are particularly detrimental to older users and thus should be avoided. In the following, age related changes to motor control are reviewed in greater detail (compared with Section 2.2.3), and, based on these findings, hypotheses are generated (abbreviated **H1-8**).

5.1 Motor Control and Aging

Age-related changes in the performance of fine motor movements have gathered much interest in occupational psychology and ergonomics already during the first decades of the 20th century (Krampe, 2002). Later, largely influenced by Welford's seminal work of the late 50s (Welford, 1958), psychological theories of aging and motor control evolved. "The general picture emerging from Welford's work and numerous studies since is that older adults' motor performance is markedly slower and less accurate than that of young individuals" (Krampe, 2002, p. 770).

Reduced movement speed. A fundamental finding across multiple studies is that older adults are generally slower in performing movements, leading to increased movement durations. In their review, Ketcham and Stelmach (2004) collected evidence for increased movement duration across a wide variety of tasks, including point-to-point movements, reaching and grasping movements, handwriting, and other continuous movements. Movement time is an important measure of how the motor system is performing. However, the observed age-related slowing effects are heavily influenced by task demands (Ketcham & Stelmach, 2004). Common para-

digms where an age-related slowing in movement can be measured include tapping tasks (decrease in the maximum repetition rate), several variants of Fitts' (1954) aiming task (e.g. Ketcham, Seidler, Van Gemmert, & Stelmach, 2002), and pursuit manual tracking (e.g. Jagacinski, Liao, & Fayyad, 1995). In general, movement durations are reported to be between 30-60 percent slower in older adults compared with younger adults, and the slowing factor increases with increasing task complexity (Ketcham, et al., 2002; Welford, 1977). Fisk and colleagues (2004) suggested, as a rule of thumb, that the movement speed of older adults can be estimated to be 1.5 to 2 times slower than that of younger adults. The slowing effects in movement manifest themselves for example in lower peak velocity amplitudes, longer deceleration profiles, and a greater need for corrective submovements in older adults as compared with younger individuals (Ketcham, et al., 2002). Reduced movement speeds of older adults were also evident in various tasks of technology interaction and on a variety of different input devices, such as keyboard, mouse, light pen, rotary encoder or touchscreen (Charness, et al., 1995; Rogers, et al., 2005; M. W. Smith, et al., 1999). In line with these studies, we assumed that older users would also perform slower at the execution of gestures than their younger counterparts (**H1**).

Reduced movement accuracy. A series of studies have documented that older adults' movements are not only slower, but also less precise. For example, the accuracy in Fitts' aiming task variants is decreasing with age (Welford, Norris, & Shock, 1969). Besides an increase in reaction times for different tasks, Fozard and colleagues also found an age-related increase in errors in both their cross-sectional and longitudinal analyses (Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994). A number of factors have been hypothesized to account for the age-related differences in movement control, including (1) higher levels of motor noise, (2) less efficient perceptual feedback, (3) different strategies of older adults, and (4) their inability to produce sufficient force for rapid movements (N. Walker, Philbin, & Fisk, 1997). Motor noise can be understood as a random error arising during the transmission of signals to muscles responsible for muscle control. The noise increases with the amplitude of movement, and, at a given force, the motor noise is greater with older adults (Chaparro, Bohan, Fernandez, Choi, & Kattel, 1999). This elevated motor noise is also considered by Liao, Jagacinski and Greenberg (1997) as the primary cause for the less smooth movement patterns (=reduced submovement accuracy) they observed in older compared with younger adults. Other findings similarly suggest that "older adults are unable to effectively propel their limb to the target in a single step, which results in multiple secondary, corrective submovements, causing their limb movement to be less smooth and slower" (Ketcham, et al., 2002). Furthermore, Sarlegna (2006) brought evidence that movement trajectories are less accurately performed by older adults because of their impairment to monitor online movement through visual feedback processes. Even in more applied tasks, e.g. when older users are required to

select adjacent buttons on a touchscreen, effects of manual dexterity on speed and accuracy could be determined (Jin, Plocher, & Kiff, 2007). Riviere, Nitish and Thakor (1996) showed that with advanced age, mouse use becomes increasingly inaccurate and nonlinear. This literature suggests that also during the execution of gesture patterns, older users' movements could be less smooth, show more corrective submovements, and ultimately perform a less accurate reproduction (**H2**) of the ideal gesture trajectory. However, there is also evidence that older adults have a bias for accuracy at the expense of speed (e.g. Salthouse, 1985; Welford, 1981). Ketcham and colleagues (2004) thus argue that in any study showing speed differences in cross-sectional age-group comparisons, also the speed versus accuracy relationship should be determined.

Reduced movement coordination. Coordination refers to the ability to control separate movement segments, or even several body parts that execute the movements, so that they match in time and space. The ability to coordinate movements has been demonstrated to become increasingly difficult with advancing age across a variety of tasks, such as aiming, reaching and grasping, drawing, handwriting, and tasks where bimanual coordination is needed (Ketcham & Stelmach, 2004). Particularly in tasks that require a complex orchestrated interplay of several joints, agonist and antagonist muscles, such as drawing or handwriting, older adults exhibit less smooth and decoupled movements, which are also slower and more variable (Ketcham & Stelmach, 2004). Since these movements are in their nature similar to the drawing of gesture patterns on a touch sensitive surface, we expect similar performances. In particular, gestures that have a higher coordination demand – gestures whose patterns are more complex and consist of a greater number of individual segments – should be more detrimental to older adults with respect to speed (**H3a**) and accuracy (**H3b**) than simpler patterns. Moreover, since several studies have documented increasing difficulties of older adults for bimanual coordination (e.g. Greene & Williams, 1996; Swinnen, 1998), we expect bimanual gesture execution to be more difficult for older users than for younger users (**H4**).

5.2 Display Size and Device Operation Posture

Besides the complexity of the gesture patterns, further external factors are hypothesized to influence the gesture execution performance of elderly users. In particular, the size of the available space to perform a gesture, and the way a device is held in the user's hand could affect how smooth and faultless a gesture pattern can be drawn.

5.2.1 Size of the Gesture Space

This research focuses on mobile devices where multi-touch screens are most abundant, and in particular on palm-sized devices such as PDAs, mobile phones and smartphones, with typical screen sizes between 3 and 4 inches. Display size has been shown to influence task performance, and to interact with task difficulty. For example, Oehl, Sutter and Ziefle (2007) tested performance on multidirectional pointing tasks on a touchscreen with three screen sizes from 3.9 inches (smartphone size) to 11.8 inches (tablet size). They found significant effects of display size and task difficulty on movement time and errors, as well as a significant interaction between display size and task difficulty. Performance generally increased with increasing display size and with decreasing task difficulty. Similar effects have also been reported in a mouse positioning task, in which positioning was found to be more time consuming for smaller displays (Tränkle & Deutschmann, 1991). If it holds true that homing, aiming and positioning tasks are in principle comparable to gesture input, we can assume that gesture performance also increases with increasing display size (H5). Furthermore, it is hypothesized that a small space for gesture execution is especially detrimental to older users. There are many reports of older people facing difficulties with buttons that are too small for error-free interaction (see Section 2.5.1). Also for movement tasks, an interaction between age and target size was demonstrated, where older adults showed differentially lower performance than younger adults as target size decreased (Ketcham, et al., 2002). Given the age-related decline in fine motor control and coordination, especially in difficult tasks (Ketcham & Stelmach, 2004), we expect older adults to face comparably more problems to accurately produce a gesture than younger users as the available drawing space decreases (H6).

5.2.2 Operation Posture of the Device

Little is known about how older users hold mobile devices in their hands, and how different postures affect interaction performance. Some indications, however, exist that older users might operate the mobile device in different positions and with different posture than younger users. For example, Siek, Rogers and Connelly (2005) observed that older users (but not the younger ones) tilted a PDA to view icons with less glare, which is in accordance with older people's increased glare sensitivity (Schieber & Kline, 1994). Furthermore, Siek and colleagues found differences in how the two groups held the PDA: older participants held the devices mainly with two hands, while younger participants held it with one hand. They attributed this observation to differences in grip strength, since high correlations between grip strength and age are documented (Mathiowetz, et al., 1985). Studies that investigated the

operation posture of mobile phones identified a general desire for one-handed interaction: the majority of users performed one-handed interaction whenever the interface *allows* for it, and fell back to two-handed interaction only when the interface *prevents* one-handed input (Karlson, Bederson, & Contreras-Vidal, 2006a, 2006b). Nonetheless, one-handed thumb-based operation brings considerable limitations, since the thumb length and rotation angle is rather small (Hirotaka, 2003), resulting in great difficulties to move in particular directions (e.g. NW $\Leftrightarrow$ SE movement), and reach certain screen regions (e.g. lower right corner, for right-handed people), due to the increased thumb flexion demands (Karlson, et al., 2006b). One-handed, thumb based gesture execution is therefore assumed to be more difficult and less accurate than two handed, index-finger based gesture execution (**H7**). Taking into account that older adults have a reduced movement range of the thumb (Allander, Björnsson, Ólafsson, Sigfússon, & Thorsteinsson, 1974; Chaparro, et al., 2000), and general problems with fine motor control and coordination (Ketcham & Stelmach, 2004), there is reason to believe that older adults have in fact also greater difficulties in one-handed mobile phone operation than in two-handed operation (**H8**). It can thus be assumed that the operation posture of the device influences how well gesture can be performed, and that operation posture affects gesture performance differentially for younger and older users.

5.3 Hypotheses

No previous studies exist that investigate gesture input performance of older adults. However, from the large body of literature concerning motor control in older adults, the following hypotheses have been derived (see previous section), which guide the empirical investigation:

- **H1: Older adults perform gestures slower than younger adults.** A main effect of age on gesture execution speed is expected.
- **H2: Older adults perform gestures less accurately than younger adults.** A main effect of age on gesture execution accuracy is expected.
- **H3: Older adults have greater difficulties in performing more complex gestures.**
 - **H3a: Older adults perform more complex gesture slower.** An interaction of age and gesture complexity is expected with respect to gesture execution time.
 - **H3b: Older adults perform more complex gestures less accurately.** An interaction of age and gesture complexity is expected with respect to gesture accuracy.

- **H4: Older users have greater difficulties in performing bimanual gestures.** An interaction of age and the number of hands required for gesture execution on gesture accuracy is expected.
- **H5: Gesture performance increases with increasing display size.** A main effect of size on gesture accuracy is expected.
- **H6: Older adults have greater difficulties in performing gestures with decreasing screen space.** An interaction of age and display size on the gesture performance accuracy is expected.
- **H7: One-handed, thumb-based gesture execution is more difficult than two-handed, index-finger based gesture execution.**
 - **H7a: One-handed gesture performance is slower.** A main effect of operation posture on gesture execution time is expected.
 - **H7b: One-handed gesture performance is less accurate.** A main effect of operation posture on gesture accuracy is expected.
- **H8: Older users have greater difficulties to perform gestures in one-handed operation mode.** An interaction of age and operation posture is expected with respect to gesture execution time and gesture accuracy.

Hypotheses **H1-H3** were tested in two studies in order to investigate whether they can be supported for stationary (study 1A) as well as mobile (study 1B) touchscreen devices, or whether the hypothesized effects are possibly restricted to a certain interface class. Furthermore, the hypotheses concerning gesturing space (**H5 & H6**) could only be investigated in a setup where space could easily be manipulated (study 1A), while the hypotheses related to the number of hands and operation posture (**H4,H7,H8**) could only be tested in the context of a mobile device (study 1B).

5.4 Methods Study 1A

The study was devised to find out to which degree the demands on motor control and coordination afforded by the gesture paradigm match the abilities of older adults, and which boundary conditions might influence the gesture performance. Gesture complexity and available screen size were systematically modified to this end. Furthermore, the subjective evaluation of different gesture patterns according to their perceived ease of use was of interest. Ultimately, implications for system design should be drawn, regarding, for example, which gesture patterns are more suitable than others for older users, which levels of complexity are still manageable, what are minimum requirements to screen size, and how detection thresholds and tolerance levels should be adjusted.

5.4.1 The Gesture Set

For the first study, a set of 42 simple one-finger gestures was devised. Since the testing device did not support multi-touch, only such gestures that could be traced with one finger were included. The gesture set was developed along the following goals. First, gesture patterns should exhibit an overall simplicity, which ensured their potential applicability in real contexts. At the same time, the patterns should allow to derive certain geometrical properties, such as angles or curvature, for the assessment of performance accuracy. Simple geometric forms (line, cross, triangle, circle) and very basic signs (cross, check mark, “scratch-out”) were chosen for this purpose. Second, comparability with existing gesture catalogues should be established. The gesture set draws reference from Microsoft Application Gestures (Microsoft, 2009), Apple’s Trackpad Gestures (Apple Inc., 2009), the GestureWorks open source gesture library (GestureWorks, 2011), and mouse gestures used for browser navigation (e.g. Opera Software ASA, 2011). Third, there was a need for systematization with respect to gesture orientation and complexity. The same gesture patterns were included in different orientations to minimize any bias of direction and orientation. Furthermore, gestures were selected so that three equally sized classes of varying gesture complexity could be established. A very basic manipulation of gesture complexity was to group gestures according to the number of line segments they contain: single-line gestures, double-lines gesture, and multi-line (3 or 4 lines) gestures. Each class consisted of fourteen gestures, which are depicted in **Figure 5-1**.

Figure 5-1. Set of finger gestures used in study 1A: single-line gestures (left), double-line gestures (middle) and multi-line gestures (right)

5.4.2 Participants and Apparatus

In total, 36 participants volunteered to take part in this study. The younger participants (9 male, 9 female) were 21-33 years old ($M=26$ years, $SD=3.4$ years), while the older group (9 male, 9 female) were aged between 60 and 71 years ($M=63$ years, $SD=3.2$ years). All participants were right-handed, had normal or corrected-to-

normal vision and no pathological condition that diminished arm, hand or finger movements, according to self-report. For their participation in this study, they received a small financial reward.

Participants were seated in front of a 15" touchscreen (Eizo L353T-C), which was mounted on a supporting wooden rack at an oblique angle of 30° for better ergonomic handling, and positioned on a desk in front of the participants (**Figure 5-2**). Gesture pattern presentation and finger input recording was realized through a custom Matlab script (MathWorks, 2011), making use of the Psychophysics toolbox (Version 3, Psychtoolbox.org, 2011). The finger movements on the touchscreen surface were sampled at a frequency of 50 Hz. The screen resolution was 1024 x 768 pixels.

Figure 5-2. Experimental setup: Presented gesture patterns (left) had to be retraced on the white drawing area (right). The largest of three possible sizes is depicted

5.4.3 Procedure

The task was to retrace visually presented gesture patterns on the touch surface. The gestures were depicted as arrows and presented centrally on the screen at different sizes (**Figure 5-2, left**). Each gesture pattern was displayed for 1.5 seconds, and afterwards an equally sized white rectangular drawing space appeared on the screen on which the gesture had to be retraced (**Figure 5-2, right**). The participants were instructed to replicate the presented gesture within the boundaries of the designated drawing area as fast and accurately as possible. They received no visual feedback of the gesture trajectory that their finger had taken on the screen in order to record the original patterns without any corrective movements.

The experimental session started with the briefing of participants, as well as a short structured interview to collect demographic data. Next, participants received a very short training session to familiarize them with the basic procedure and the task

demands. Six gestures had to be drawn, two in each size. The training gestures were selected to include many of the features of the total gesture set, including a straight and a curved gesture, a single-line, a double-line and a multi-line gesture, and a gesture consisting of several segments (*cross* gesture).

After the training, the participants had to complete nine blocks of gesture tracing. Upon completion, they received a questionnaire to rate the task difficulty. Mental Demand and Physical demand subscales from the NASA Task Load Index (Hart & Staveland, 1988) were assessed, and participants were asked to single out particularly easy and particularly difficult gestures. Furthermore, participants had to state which *features* of gestures they perceived as difficult, and were asked about previous as well as intended use of gesture controlled devices. Before the session was concluded, participants had to fill out questionnaires regarding their technology affinity (Karrer, Glaser, Clemens, & Bruder, 2009) and their general action strategies, such as the individual trade-off between speed and accuracy (Nachtwei, 2008a, 2008b). The whole session lasted approximately one hour.

5.4.4 Design

The quasi-experiment was conducted using a 2 (AGE GROUP) $\times$ 3 (SIZE) $\times$ 3 (COMPLEXITY) factorial design. AGE GROUP varied as between-subject factor (younger users, older users), while SIZE (small, medium, large) and COMPLEXITY (single line, double line, multiple line) varied as within-subject variables. The size of the areas was chosen so that the medium size corresponded approximately to the size of an iPhone display (reference size), while the small size covered $\frac{1}{4}$ of that area (corresponding to sub-palm sized devices, lower bound), and the large size 4 times the area (corresponding to tablet or stationary touchscreen). The dimensions of the designated gesture drawing areas on screen were:

- Small: 2.6 x 3.8 cm, corresponding to a 1.8 inch display
- Medium: 5.2 x 7.5 cm, corresponding to a 3.6 inch display
- Large: 10.4 x 15.1 cm, corresponding to a 7.2 inch display

Each of the 42 different gestures was repeated three times per size, resulting in a total of $42 \times 3 \times 3 = 378$ trials, organized in 9 blocks. The repetition of individual gestures minimized the risk of missing data, and since averaged data was used in subsequent analyses, increased reliability of the measurement. Within each block, the size of the gesture area was held constant, but the order of gestures was randomized. Each block contained 14 gestures from each complexity level (**Figure 5-1**). Between blocks, participants were allowed to pause if needed. The sequence of blocks was permuted across participants, so that each size appeared equally often at each serial position,

and that the younger and older participants received the same sequences at equal frequencies.

5.4.5 Data Analysis

The performance of gesture execution was measured by a number of different parameters. *Error rate* (when the gesture was not retraced correctly) and *boundary violations* (when the finger passed beyond the designated drawing space) provided measures of task fulfillment. An error occurred whenever the basic gesture shape could not be found in the recorded movement pattern. In particular, movement patterns with too few sample points, where the finger did not move at all, or where the number of segments did not match the gesture template (e.g. in a two-line gesture only one line was drawn), were regarded as errors. The error measure was quantified as percentage of the total amount of performed gestures. Each movement pattern that exceeded the drawing area by more than 1mm (4 pixels) was counted as a boundary violation. The *average velocity* was calculated between the first touch-down and the last touch-up event. The measure of gesture accuracy was divided in parameters of *form stability* and parameters of *directional stability*.

Form stability provides a rough estimate to which degree the original gesture shape was preserved in the gesture movement pattern produced by the participant. Importantly, the measure of basic shape preservation had to be size and position invariant. This parameter had to be assessed differentially for linear and circular gestures. For linear gestures (all gestures that consisted of a straight line, or a sequence of straight lines), form stability was measured by the *angular deviation* from a reference angle. To this end, straight lines were fitted to each linear segment of a gesture through an orthogonal least square fit, and the angle between the linear fit and a reference line was determined (**Figure 5-3, left**).

Figure 5-3: Gesture accuracy measurements: linear fit and angle measures (left), ellipse and circle fit (right)

For double- and multi-line gestures, the angles between the linear fits of the separate segments were measured, compared with reference angles and the deviations were averaged. For circular gestures, form stability was measured by the *eccentricity* value of the best ellipse fit to the data. Eccentricity can be understood as the degree of roundness of the pattern, i.e. the degree to which the perfect circle from the gesture template could be retained. A value of 0 denotes a perfect circle, and increasing values denote increasing asymmetry of the ellipse (**Figure 5-3, right**). The ellipse fitting procedure was performed (using the *fitellipse.m* script by R. Brown, 2007) by nonlinear least squares, optimizing the squared sum of orthogonal distances from the points to the fitted ellipse.

Directional stability was established as a measure of how smoothly the pattern was executed, or how much shakiness was observable in the movement trajectory. Again, this parameter had to be assessed in slightly different ways for linear and circular gestures. For linear gestures, directional stability was simply measured as the average deviation from the movement pattern to the best orthogonal least squares linear fit for each segment. For circular gesture patterns, the average distance to the best geometrical ellipse fit was calculated instead. The measures of directional stability were labeled *line deviation* and *ellipse deviation* respectively.

All measures of velocity, form- and directional stability were analyzed only on non-erroneous gestures. For each dependent variable, a repeated measures analysis of variance (ANOVA) was calculated, with the between-subjects factor AGE GROUP, and the within-subjects factors SIZE and COMPLEXITY. Several ANOVAS had to be calculated instead of a MANOVA, since not all cells of the factorial design were covered by each dependent variable. For example, eccentricity and ellipse deviation could only be measured for curved gestures, which existed only for complexity level one. For some analyses, the assumption of sphericity was not met, and the degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity.

5.5 Results Study 1A

5.5.1 Sample Description

Both the younger ($M=26$ years, $SD=3.4$ years) and older group ($M=63$ years, $SD=3.2$ years) showed relatively high levels of education. In the young sample, over 90% were students or had already graduated. In the older sample, two thirds of the participants had received higher levels of education. Both groups had already some previous experience with mobile phones: all participants except one younger subject

stated they currently use a mobile phone, their average years of usage was between 9.7 years (older) and 6.3 years (younger), and the average number of short messages written per month differed ($U=81.5$, $z=-2.56$, $p=.009$, $r=-.43$) between older adults ($Mdn=3$) and the younger group ($Mdn=25$). All participants in both groups reported to have used touchscreens before, e.g. in ATMs or ticket machines. When asked whether they had controlled a device by the use of gesture before, 83% of the older and 56% of the younger group gave a negative response, but this difference was not significant ($\chi^2(1)=3.27$, $p=.146$). Furthermore, both groups were comparable regarding their technology affinity (as measured with the TA-EG instrument, Karrer, et al., 2009). None of the subscales (enthusiasm, competence, negative effects and positive effects) showed any significant difference between the two age groups. Also with respect to their action strategy (Nachtwei, 2008b), no group differences were observed. This means, contrary to previous reports, the older participants in this study did not show a different speed-accuracy bias than the younger group ($F(1,34)=3.19$, $p=.083$, $\eta^2=.086$). Differences found in gesture execution speed and accuracy can thus be expected to arise more likely from age-related differences in motor control rather than different strategies that favor accuracy over speed or vice versa.

5.5.2 Gesture Performances Measures

A summary of descriptive results can be found in **Tables 5-1** and **5-2**. The separate parameters will be discussed in more detail in the following.

Table 5-1. Comparison of dependent variables according to *age group*.
Asterisks mark significant main effects (** $p<.01$, * $p<.05$)

dependent variable		younger		older	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
error rate	(%)	2.65	3.08	4.54	5.87
boundary violation rate	(%)	4.50	6.00	5.00	5.79
average velocity	(cm/s)	13.91	4.37	10.66	3.70 *
angular deviation	(°deg)	5.39	0.66	5.46	1.21
eccentricity		0.60	0.04	0.62	0.07
line deviation	(mm)	0.69	0.15	0.68	0.14
ellipse deviation	(mm)	0.50	0.14	0.53	0.23

Table 5-2. Comparison of dependent variables according to different levels of *size* (small, medium, large) and *complexity* (single-line, double-line, multi-line).

Asterisks mark significant main effects (** $p < .01$, * $p < .05$)

size:		small		medium		large		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
error rate	(%)	6.09	7.19	3.30	5.18	1.39	2.77	**
boundary violation rate	(%)	9.14	12.48	3.62	5.59	1.48	2.07	**
average velocity	(cm/s)	6.45	2.61	11.43	3.82	18.97	5.99	**
angular deviation	(°deg)	6.53	1.09	5.02	1.03	3.93	0.85	**
eccentricity		0.65	0.06	0.61	0.06	0.57	0.07	**
line deviation	(mm)	0.46	0.10	0.65	0.15	0.94	0.22	**
ellipse deviation	(mm)	0.30	0.10	0.49	0.23	0.76	0.23	**

complexity:		single		double		multi		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
error rate	(%)	3.27	4.82	3.72	4.76	3.79	4.48	
boundary violation rate	(%)	4.60	5.83	4.65	6.36	4.99	6.17	
average velocity	(cm/s)	12.93	4.60	12.99	4.26	10.93	3.38	**
angular deviation	(°deg)	3.42	0.66	6.43	1.25	5.62	1.60	**
eccentricity		0.61	0.06					
line deviation	(mm)	0.67	0.14	0.70	0.17	0.68	0.16	
ellipse deviation	(mm)	0.52	0.18					

Error rate and boundary violations. Both parameters of task fulfillment showed no significant effects of age group or complexity. Older users performed an erroneous gesture in 4.54% of cases ($SD=5.87$), whereas younger users erred at only half this rate ($M=2.65\%$, $SD=3.08\%$). However, due to huge individual differences, no significant group effect could be observed. The designated gesture space was left at comparably equal rates (5% for the older users, 4.5% for the younger users). No effects of gesture complexity on the error rate or on the boundary violation rate could be observed, nor were there significant interactions between complexity and age group. There was, however, a stable influence of size (=the available space to draw the gesture) on error rate ($F(1.4,48.5)=12.94$, $p < .001$, $\eta^2=.276$) and boundary violation rate, $F(1.3,43.1)=12.96$, $p < .001$, $\eta^2=.276$). The larger the available size, the fewer errors were made (**Figure 5-4 left**), and logically, the less frequently did users hit the boundaries (**Figure 5-4 right**). This effect was similar for older and younger users alike, and no interaction of size and age group could be observed.

Figure 5-5: *left*: Mean error rate according to available gesture space; *right*: Mean percentage of gesture space violations for three different sizes; Error bars denote a 95% confidence interval.

Average velocity. As predicted by hypothesis **H1**, older adults performed the gestures significantly slower ($M=10.7$ cm/s, $SD=3.7$ cm/s) than younger adults ($M=13.9$ cm/s, $SD=4.4$ cm/s) across all sizes and complexities, $F(1,34)=6.671$, $p=.014$, $\eta^2=.164$. In addition, there was a significant effect of size on gesture execution speed: the larger the gesture to retrace, the faster it was performed, $F(1.1, 37.9)=340.3$, $p<.001$, $\eta^2=.909$, **Figure 5-5, right**). Repeated contrast analysis revealed that the medium size was significantly faster than the small size ($p<.001$), and also the large size was faster than the medium size ($p<.001$). Furthermore, the complexity of the gesture pattern influenced how fast it was drawn, $F(1.4, 46.8)=52.38$, $p<.001$, $\eta^2=.606$ (**Figure 5-5, left**). Contrast analysis revealed that while the single-line and double-line patterns did not differ, the multi-line condition accounted for a significant slowing. Apart from these main effects, we found an interaction of age group and display size ($F(1.1,37.9)=12.139$, $p=.001$, $\eta^2=.263$): even though both age groups increased their gesture drawing speed with larger sizes, the younger users could take more advantage of the largest size. No interaction between age group and complexity as hypothesized (**H3a**) was found.

Figure 5-4. *left*: Gesture performance velocity according to different levels of gesture complexity; *right*: Gesture performance velocity according to available gesture space; Error bars denote a 95% confidence interval.

Form stability measures. Form stability was measured through angular deviation for linear gestures and eccentricity for circular gestures. Angular deviation was influenced by size ($F(2,68)=232.5$, $p<.001$, $\eta^2=.872$) and complexity ($F(1.7, 56.6)=83.65$, $p<.001$, $\eta^2=.711$), but no significant differences due to age group could be observed. The larger the gesture pattern, the more closely could participants of both groups reproduce the original angles, and thus provide a better replication of the prototypical pattern. The average deviation from the correct angle dropped from 6.5° in the smallest size, over 5° in the medium size to 3.9° in the largest size (**Figure 5-6, left**). A similar result was obtained for the eccentricity measures of the best fitted ellipse: a strong influence of size ($F(2,68)=60.69$, $p<.001$, $\eta^2=.641$) was observed, but no influence of age group (**Figure 5-6, right**). Since circular gestures did only exist as single-line gestures, no influence of complexity could be determined. For both form stability measures, contrast analysis revealed a highly significant linear trend for the factor size ($F_{\text{ang}}(1,34)=354.33$, $p<.001$, $\eta^2=.912$; $F_{\text{ecc}}(1,34)=121.18$, $p<.001$, $\eta^2=.781$). Pairwise comparisons (Bonferroni corrected) between the small, medium, and large conditions showed that mean values of form stability differed significantly across all comparisons. Comparison of the angular deviation between the single, double and multi-line gestures revealed that all three complexity levels differed significantly from another, with the smallest accuracy observed in the double-line condition. Why the gestures in this category seemed to have been especially difficult to reproduce is discussed below (Section 5.6.1). Taken together, the measures of form stability show a strong influence of display size on the accuracy of gesture execution (**H5**), but no sign of reduced accuracy in the older user group (**H2**, **H3b**).

Figure 5-6. *left*: Average angular deviation of gesture patterns according to display size; *right*: Eccentricity of fitted ellipse according to display size; Error bars denote a 95% confidence interval.

Direction stability measures. For both circular and linear gestures, a main effect of gesture size on the smoothness of the gesture pattern was observed: with increasing

gesture space, there were significantly larger deviations from the best linear fit $F_{\text{line}}(1.4, 46.4)=232.64, p<.001, \eta^2=.872$) or best ellipse fit $F_{\text{ellipse}}(2,68)=155.44, p<.001, \eta^2=.821$ respectively. In both cases, repeated contrasts showed that the largest size produced larger deviations than the medium size, which in turn revealed larger deviations than the smallest size. While these results might be counterintuitive at first glance, it becomes clear that the larger the gesture movement itself, the larger was also the jittering of the recorded trajectory around the ideal path. No main effect of complexity was found. Moreover, there was again no main effect of age group for both measures, indicating that across all gestures and sizes, older and younger users did not differ in how smoothly they produced their gesture trajectories (**H2**). However, there was an interaction of age group with size regarding the line deviation measure, $F(1.4,46.4)=5.536, p=.015, \eta^2=.140$), with significant differences between the small and medium ($p=.043$) as well as the small and large ($p=.014$) conditions. While in the smallest gesture space the younger participants were more accurate, both groups were roughly equal at the medium size, but the older group was superior in terms of direction stability at the largest size. A significant three-way interaction between age group, complexity and size ($F(4,136)=5.46, p<.001, \eta^2=.138$) on this measure further differentiates the previous interaction. Contrast and pairwise comparison analyses showed that the accuracy advantage on small screen sizes for younger users was independent of gesture complexity, but with larger screen sizes, also the influence of complexity became stronger. It can be seen that the accuracy benefit at the largest screen size for older users is modulated by complexity, such that the more complex the gestures get, the bigger is the direction stability difference between older and younger users. Taken together, these results could not support a general decline in accuracy of older users (**H2**), nor could they prove particular difficulties with complex gestures (**H3b**). Some indications for **H6** (older users have greater difficulties at decreasing screen sizes) were obtained.

5.5.3 Subjective Gesture Evaluation

After the gesture performance procedure, participants were asked to evaluate the gesture task with regard to their perceived mental and physical demand, using the respective subscales of the NASA-TLX questionnaire. The scales had 21 levels, from 0 (very low demand) to 20 (very high demand). Mental demand was rated low both among older ($Mdn=5$) and younger ($Mdn=5$) participants, and no age differentiation was observed ($U=149.0, z=-0.133, p=.90$). Physical demand was also perceived low in general, but was rated slightly higher ($Mdn=8$) in the younger group than in the older group ($Mdn=3; U=93.0, z=-1.997, p=.046, r=-.34$). One reason for this somewhat counterintuitive result might have been that younger participants rated the physical effort not absolutely, but in comparison with their touchscreen experiences of con-

temporary mobile devices, which mostly operate with less touch pressure than the touchscreen device used in the test.

Furthermore, participants were asked to select five gestures which they perceived as particularly easy, and five which they perceived particularly difficult. Based on the relative frequency of being selected as easy or difficult within each age group, a rank order was established, of which the ten best and worst ranked gestures are depicted in **Figure 5-7**. Straight, single-line gestures were perceived as easiest to perform. The up-down direction furthermore seems to be easier than the right-left direction, and within up-down, the down direction seems to be perceived as easier. This could be an effect of *pulling* a finger towards you is perceived easier as *pushing* the finger away from you. As most difficult appear gestures with multiple lines, since the two 4-line gestures of the gesture set were rated most difficult. Furthermore, within the three-line gestures, those which do not have right angles are perceived as more difficult. Interestingly, also circular patterns, even though grouped into the single-line complexity class, were perceived as difficult by the users. The subjective difficulty rating of a gesture correlated only moderately with accuracy measures (e.g. with angular deviation $r_{old} = -.27$, $r_{young} = -.36$; or eccentricity: $r_{old} = .46$, $r_{young} = .44$).

Figure 5-7. Gestures rank ordered according to their average perceived easiness/difficulty. For both age groups, the differential (easy-difficult) relative frequencies for the top and bottom ten gestures are displayed.

As can also be seen in **Figure 5-7**, subjective evaluation of difficulty was highly similar between younger and older users ($r_s=.72$, $p<.001$). Looking at the gesture rankings with the largest differences between older and younger users, we find that *flash_right*, *plus* and *check_mark* are the gestures that had the strongest advantage among older users compared with the younger group. Likewise, *circle_clockwise*, *mirror_plus* and *semicircle_up* showed the largest difference in favor of the younger group. Even though no quantification of different features is possible, the following observations can be made: older users seemed to rate such gestures easier that have a strong iconicity and semantics (gestures that are meaningful symbols), which was not observed for the younger users. Contrarily, since two of the gestures ranked notably better by the younger users are circular in nature, this type of gesture patterns might be perceived as relatively easier among the younger users, or more difficult among the older ones. This observation is backed up with subjective descriptions of which features were perceived particularly difficult. Older users reported for example difficulties in “*exactly retracing the curvatures*”, they complained about the required “*marksmanship for circles*”, or the difficulty to make a “*precise distinction between full- and semicircles during performance*”. Four of the 18 older users mentioned they perceived it easier to draw gestures downwards than upwards, and three older users mentioned a bias for gestures performed from left to right, “... *since one is used to write from left to right*”. Several older users also mentioned an increase in difficulty with the number of corners and direction changes, three of which particularly mentioning the *flash* (zigzag) gestures as particularly difficult (“*greatest source of errors*”). The younger users commented in similar ways: they also found that “*upwards (away from the body) and leftwards stroking was very cumbersome*”, they deplored frequent changes of direction, especially at acute angles, and claimed difficulties in drawing exact circles and semicircles. Finally, when asked about their willingness to use a technical device “with the help of the tested or similar gestures” in the future, there was no difference between the younger and older users’ responses ($U=141.50$, $z=-.713$, $p=.59$, $r=-.12$). Within the older user group, 28% strongly agreed (younger: 22%), 55% rather agreed (younger: 50%), and 17% rather disagreed (younger: 28%). No participant stated a strong disagreement.

5.6 Discussion Study 1A

5.6.1 The Influence of Space and Complexity on Gesture Performance

Across all participants of the study, it was observed that the space available to draw a gesture strongly influenced the way the gesture was performed (see **Table 5-2**). In

the case of errors and boundary violations, the rates of unsuccessful gesture execution rise from acceptable 1.4-1.5% on the largest screen to 6-9% on the smallest screen size. This should, however, raise not too much concern, since the smallest size was deliberately chosen as a lower bound estimate, and hardly any devices (except the latest iPod shuffle) exist that employ multi-touch screens of such small dimensions. Moreover, the boundary in this study was realized only through a visual frame, thus the generalizability to physical screens of comparable sizes is not necessarily given. In fact, it seems rather unlikely that users would overshoot gestures on a physically bounded screen to the same extent, but rather trim and squeeze the gesture patterns to fit the screen area.

The variation of gesture space had also a strong effect on the gesture performance speed: the larger the pattern that had to be retraced, the greater the velocity. The participants were about twice as fast for the medium sized gesture patterns, and three times as fast for the large patterns compared with the smallest size. This means, while with a quadratic increase in size, the execution speed increased roughly linearly. From a psychological point of view, this observation is in line with kinematic laws that describe basic principles of human motor pattern generation. It reflects the principle of *isochrony*, (Freeman 1914, cited from Gibet, Kamp, & Poirier, 2004), stating that there exists a spontaneous tendency to increase the velocity of motion according to the distance to run, when no constraint on mean velocity is imposed. Freeman analyzed writing movements, which are in nature rather similar to the drawing of gesture patterns on a touchscreen. A similar psychophysical law, which is referred to as "*Gesetz der konstanten Figurzeit*" (Derwort, 1938, cited from Heuer, 1984, translation: "law of constant figure tracing time") postulates that the time needed to trace identical figures of different size is constant. From a pragmatic point of view, the velocity of gesture execution is only of secondary importance. It does not matter so much whether the user performs a gesture faster or slower, as long as its execution is not too careless and accuracy is preserved to a point where it can be correctly classified by the device.

With regard to the observed accuracy of gesture execution, the data showed that the measures of form stability and directional stability have been affected differentially by the size: with increasing drawing space, the basic form was reproduced more stably (as seen in reduced angular deviation and smaller eccentricity values of the fitted ellipse), while directional stability (as seen in the increase deviation from the best linear or ellipse fit) increased. This finding might be surprising at first glance, but can be explained in the light of a speed-accuracy tradeoff. Directional stability measures the degree to which the drawn patterns deviate from the optimal path, and thus accumulates all overshoots and readjustments. Overshooting, slipping and readjusting are more common – and have greater effects – at higher velocities, which

were observed with increasing gesture size. Therefore, a drop in directional stability with increasing size of the gesture patterns is understandable. On the contrary, the basic shape of the pattern is better preserved at larger sizes. In the case of angular deviations, this becomes obvious from the fact that the same small displacement of the finger on the screen can result in large angle deviations if the reference pattern itself is small, but results in smaller angle deviations if the reference pattern is larger (e.g. an end point displacement by 1mm results in 11° angle deviation against a 5mm reference line, but in only 1° against a 50 mm reference line). As hypothesized in **H5**, a clear effect of display size on gesture accuracy was observed. If form stability is taken as the more general description of how well the gesture is preserved – which is probably also of higher relevance for gesture detection algorithms – we can further state that gestures are redrawn with greater accuracy if more space for performing the gesture is available.

The manipulation of gesture complexity, as a simple variation of the number of lines the gesture is made up of, did not result in a systematic effect across the dependent variables. Despite a higher number of gesture segments at complexity levels two and three, participants did not slip up significantly more often, and thus no effect on errors was observed. Even though they did not make more errors, participants were significantly slower at higher complexity levels. This decrease in average velocity with an increasing number of line segments is explained well by the fact that users have to decelerate and accelerate again at each turning point of the pattern. More interesting is the effect of gesture complexity on angular deviation, the only main effect observed with regard to accuracy assessment. The highest angular deviation was found for double-line gestures followed by multi-line gestures and single-line gestures. This pattern could be explained by the geometric properties of the gestures that make up the groups: the double-line gestures consisted of many gestures which included an acute angle (see **Figure 5-1**), while the multi-line group included more rectangular angles, and the reference angles in the single-line categories were only 90° or 45°. It seems that especially right-angled patterns (but also 45°) can be replicated more accurately than patterns that include more acute angles of less common degrees. None of the other accuracy parameters showed an effect of gesture complexity. The reason for this lack of a systematic effect of gesture complexity is most likely found in the initial division into single-line, double-line and multi-line gestures, which did not reflect the actual difficulty of the gesture. It can be easily conceived that a circular gesture poses much higher demands on motor control than a single horizontal line, even though these were both part of the same category. Our classification was a simple one-dimensional approach, but the qualitative data hints at a multi-dimensional account, including for example also curvature, symmetry, direction, or familiarity of the gesture.

5.6.2 The Influence of Age Group on Gesture Performance

The central question of this study is whether there are any age-related differences in gesture execution performance, in particular with respect to available screen space and levels of complexity of the gesture pattern. Across the different measurements, the results showed generally an influence of age on the execution speed, but not on error rate or the accuracy of gesture performance. Bearing in mind the limitations of the restricted user and gesture sample, these results suggest that older users have no more problems than younger users in performing accurate gestures, even for more complex gesture patterns and even on very small screen devices.

In our gesture retracing task we found a slowing factor of 1.3, which is at the lower end of reported efficiency drops in studies concerned with age effects on input device performance. In comparison, Smith et al. (1999) reported that motor performance dropped by the factor 2.5 to 3 for middle-aged and older mouse users as compared to younger ones. Chaparro et al. (1999) quantified the age-related decline in movement speed of older mouse and trackball users by the factor of 1.6. In another study by Sutter et al. (2007), slowing factors of 1.9 (trackpoint) and 1.3-1.4 for mouse and touchpad tasks were reported. Even though varying tasks, input devices and age-bands make it hard to compare slowing effects across the studies, the slowing effects observed in the present study are within the same range, albeit at the low end. Given the rather low factor, it can be speculated that a generation gap might be less pronounced for the touch- and gesture interaction paradigm as compared other input methods (e.g. mouse).

Age differences in velocity were smallest (1.59 cm/s) for the smallest screen size and largest (5.81 cm/s) for the largest screen size (cf. **Figure 5-5**). It thus seems to be the case that the speed advantage found for younger users is mainly generated by their increased velocity in the largest screen size. It is known from aging psychology research that older adults prefer accuracy over speed (Salthouse, 1985; Welford, 1981). Evidence for this behavior has also been demonstrated in technology interaction tasks, for example in mouse and trackball usage (Chaparro, Bohan, Fernandez, Kattel, & Choi, 1999). Sülzenbrück and colleagues (2010) note that older people even favor accuracy over speed if they are instructed to perform as quickly as possible. In fact, this might be the reasons why there are no obvious drawbacks in gesture accuracy for the older adults in the present study, neither on the form stability nor on the directional stability measures.

Interestingly, in their comparison of age-related slowing effects across different motor tasks, Sülzenbrück and colleagues (2010) found task contingent slowing effects and even a performance advantage for elderly users in the line-tracing task, which resembles these gesture retracing tasks most closely. The authors explain this obser-

vation with its close relation to handwriting, in which, as has been shown previously, older adults have longer experience. While their highly controlled line tracing task was quite different in nature (following a designated groove course with a pen without touching its walls) and this might account for the divergent findings, their results are encouraging in two ways: First, if a large sample of older adults ($N=180$) can outperform a younger group on a line-tracing task, which is highly similar to performing an accurate gesture pattern on the touch screen surface, then gestures as input paradigm might not be totally unsuitable for older adults. Second, considering the fact that elderly have usually a high proficiency in handwriting, they might benefit from gesture patterns that are borrowed from or similar to patterns used in cursive writing. This assumption is backed up by the gesture ratings and subjective reports (Section 5.5.3), which suggest that symbolic gestures are more favored by the older adults.

5.7 Methods Study 1B

The second study was very similar to the first one. The basic task of retracing a previously shown gesture pattern was maintained, as were the measurement parameters of speed and accuracy. Furthermore, the gesture set was highly similar to the one used in study 1A. The major difference was to take gesture interaction from a stationary screen to a mobile device. Since mobile devices are currently the most successful application field of touch- and gesture interaction, and in the focus of the present research, the second study added external validity and relevance. Two new aspects could be investigated: First, it was now possible to assess the influence of device operation posture on gesture performance. Second, due to the multi-touch capabilities of the testing device (iPod Touch), it was possible to investigate also gestures in which more than one finger moves simultaneously. We restricted ourselves, however, to single- and two-finger gestures.

5.7.1 The Gesture Set

The gesture set used in the study 1B differed from the previous gesture set in three ways. Most importantly, it included not only single-finger gestures, but also sixteen patterns that had to be performed with two fingers simultaneously (see **Figure 5-8**). Similar to the previous gestures, the new patterns were chosen with respect to already used or suggested patterns from industry and research (e.g. the *pinch* and *spread* gestures advocated by Apple). In particular, the additional gestures consisted of circular as well as linear movement patterns. Second, to keep the total set manage-

able, only 30 of the 42 single-finger gestures from the first study were reused. While all the basic shapes were preserved, they were no longer included in all orientations. Third, since the initial segmentation in three complexity classes according to the number of line segments had been shown to be too crude as distinguishing element and no consistent effects of this complexity manipulation could be established in the first study, the gesture set was three-partitioned according to the user ratings of perceived difficulty. Of the thirty single-finger gestures included, the ten best rated gestures formed the *easy* category; the next ten gestures formed the *medium* category, while the ten worst rated gestures comprised the *difficult* category.

Figure 5-8. Gesture set used in study 1B. *Top*: Single-finger gestures according to three difficulty categories. *Bottom*: Two-finger gestures.

5.7.2 Participants and Apparatus

Thirty-nine people participated in the study, of which one older adult was later excluded because he turned out to be too old for the intended age-band (60-75). The nineteen remaining older adults (9f, 10m) were 60-74 years old ($M=66$ years, $SD=1.5$ years), while the nineteen younger adults (10f, 9m) comprised individuals of 20-32 years ($M=26$ years, $SD=3.4$ years). All participants were right-handed, had normal or corrected-to-normal vision, and except for one participant, they all reported no

pathological condition that diminished arm, hand or finger movements, according to self report. One participant reported to suffer from mild arthrosis in the thumb, but remained in the sample, as she stated that this did not impede movement capabilities in daily activities, and she could continue pursuing leisure activities such as piano playing or knitting. For their participation, participants received a small financial reward. No participant from the previous study was included in this sample.

Participants were seated at a table at a height and distance adjusted to their comfort. The device on which the gestures had to be performed was an iPod Touch (2nd generation, 8GB). The device measured (HxWxD) 110 mm x 61.8 mm x 8.5 mm and weighed 115 grams. It included a 3.5 inch multi-touch display, with a resolution of 480 x 320 pixels (163 ppi). The device was either held in the participants' preferred hand, or, in one experimental condition, placed into a fitting frame on the table, which could be repositioned to participants' comfort and was fixed by Velcro to the table to provide stability throughout the experimental block. The visual presentation of the gesture patterns as well as the movement trajectory recording and the logging of the data were accomplished through a custom-written tool developed with Xcode in Objective-C. The user input was sampled at 50 Hz.

5.7.3 Procedure

Analogous to the first study, the task was to retrace previously visually presented gesture patterns on the multi-touch surface of the device. The gesture patterns were displayed full screen and depicted by arrows as shown in **Figure 5-8**. Each pattern was presented for 1.5 seconds, and afterwards a white screen appeared on which the gesture had to be drawn. No visual feedback of the trajectory was given. Participants could proceed to the next gesture by pressing a "next" button at the bottom of the screen. Participants were instructed to perform the gesture as accurately and fast as possible.

Each session began with a short training, in which eight prototypical gestures had to be drawn so that participants got acquainted with the task and device. The training gestures contained at least one representative from each of the three difficulty categories (see Section 5.7.4), included linear as well as circular gestures and comprised single-finger as well as two-finger gestures.

The gesture performance task itself was organized in two parts. The first part consisted only of single-finger gestures, while the second part consisted only of two-finger gestures. Part one included nine blocks of 30 gestures, and part two six blocks of 16 gestures. After each block, participants were allowed to take breaks if needed.

Each participant had to retrace the gestures in three different operation posture conditions (see **Figure 5-9**). As a baseline, the device was mounted using the frame it

was shipped in on the table at a fixed position (condition “*table*”). In this condition, the device was immobilized and could be operated with one or both hands, providing the greatest flexibility and freedom of choice which fingers to use. In the second condition, participants were instructed to hold the device in their left hand (since they were all right-handed), while the gestures had to be performed with the fingers of the right hand (condition “*hand*”). In virtually all cases, the pattern was drawn with the index finger. In the third condition, participants were asked to take the device in their strong (right) hand, and perform the gesture patterns single-handedly, using only their thumb (condition “*thumb*”). This was the most restrictive operation mode, offering no degrees of freedom on how to operate the device.

Figure 5-9. Three different operation posture conditions. *Left*: Device stabilized on table. *Middle*: Hand-held, operation with 2nd hand, b. *Right*: Unimanual operation, device operated by thumb.

After the gesture performance task, participants had to rank all available gestures according to their perceived difficulty, separately for single- and multi-finger gestures. To this end, cards depicting the gesture patterns had to be sorted on a table in a sequential manner. Before the start of the actual task, demographic and prior experience data was gathered in a structured interview. After the experimental blocks were completed, questionnaire data was gathered including the Mental Demand and Physical Demand subscales from the NASA Task Load Index (Hart & Staveland, 1988), the TA-EG questionnaire on technology affinity (Karrer, et al., 2009), and the FESE-R questionnaire on action strategies (Nachtwei, 2008a, 2008b). The whole session lasted between one and 1.5 hours.

5.7.4 Design

Part one of this quasi-experiment was conducted as a 2 (AGE GROUP) $\times$ 3 (DIFFICULTY) $\times$ 3 (POSTURE) factorial design. AGE GROUP varied as between-groups factor (younger users, older users), while DIFFICULTY (easy, medium, hard) and POSTURE (table, hand, thumb) varied as within subject variables. The levels of difficulty had been derived from perceived difficulty rankings of the gesture patterns from study 1A. Three different levels of operation posture were tested (**Figure 5-9**).

Each of the 30 gestures (10 of each difficulty level) had to be performed once per block, and the order of gestures within a block was randomized. Within a block the operation posture was kept constant. Each gesture had to be repeated three times in each of the three operation posture conditions, resulting in 9 blocks and 270 gesture tasks in total. Repetitions were included to increase reliability of measurement and minimize the risk of missing data. The order of posture modes (=blocks) was permuted for each age group, such that sequence effects were minimized.

For the two-finger gestures in part two, no data on their perceived difficulty was available from study 1A, therefore these gestures were not subdivided into different categories. Furthermore, two-finger gestures could not be performed when only the thumb was available for execution, and thus these gestures were only tested in two operation posture conditions: table and hand.

Consequently, the second part was structured as a 2 (AGE GROUP) $\times$ 2 (POSTURE) factorial design, with AGE GROUP as between and POSTURE as within subject factors. The second part consisted of six blocks (3 repetitions for each posture condition) with 16 gestures each, resulting in a total of 96 gesture performances. Again, each gesture was repeated three times in each posture condition in order to obtain a more reliable data basis. The order of the posture conditions was again permuted and equalized across age groups, while the sequence of gestures within a block was randomized.

5.7.5 Data Analysis

In analogy to the previous study, the performance data was analyzed with respect to errors, speed of gesture execution and accuracy of gesture execution. The *error rate* was established as a compound measure of all gestures that were not executed correctly (or at all) as a percentage of the total number of gesture trials. Errors, in this sense, comprised gestures that were not recorded at all, that contained too few sample points (<10), that included missing segments (e.g. only two sides of a triangle are drawn), that contain additional elements, and also those gestures where the ellipse fitting algorithm did not converge. As a difference to the first study, the error catego-

ry also included gestures where only one finger was recorded for a two-finger gesture. The *average velocity*, as well as accuracy parameters of *form stability* and *directional stability* were assessed as described in Section 5.4.5. For two-finger gesture patterns, the velocity and accuracy parameters were measured independently for each finger. For the four rotary gesture patterns where one finger was supposed to rest at a fixed position while the other finger rotates around it (cf. **Figure 5-8**), only the rotating finger was measured. For the other two-finger gesture patterns the average of both fingers was taken. For each dependent variable, a mixed design analysis of variance (ANOVA) was conducted. Separate analyses were conducted for single and dual finger gestures, since the factors and factor levels differed. Similar to study 1A, multivariate analysis did not apply since not all dependent variables could be assessed in each cell of the factorial design. For example, for the measures of ellipse deviation and eccentricity, a 2 (AGE GROUP) $\times$ 2 (DIFFICULTY) $\times$ 3 (POSTURE) repeated measure ANOVA was performed, since curved gestures existed only in two of the three difficulty classes (cf. **Figure 5-8**). In cases where the assumption of sphericity was violated, the Greenhouse-Geisser correction is reported. In some cases, the assumption of normally distributed data was violated in one or both groups. In these cases, non-parametric tests (Mann-Whitney test, Wilcoxon signed-rank test, Friedman's ANOVA) were conducted for each independent variable (cf. **Appendix A.3.2**). In cases where ANOVA proved robust against the normality violation, the results of ANOVA are reported to achieve consistency. One older participant had to be excluded from the analysis of single-finger gestures, since more than half (60%) of the gestures were missing in the recorded logfile. Since no particular problems became apparent during the experimental session and since the participant could perform two-finger gestures within normal error margins, this is believed to be caused by a technical problem in the logfile creation rather than the participant's inability to perform the gestures.

5.8 Results Study 1B

The second study was very similar in design, dependent variables and results as the first one. The results section will cover therefore only the differences to the previous study and focus on the new aspects of device operation posture and multi-finger gestures. The details of the analysis can be found in **Appendix A.3**.

5.8.1 Sample Description

The sample in this study was highly similar to that in study 1A. Both showed high levels of education. While all participants in the young group had graduated or were students, also more than half of the older group had received higher levels of education. All participants had normal or corrected-to-normal vision according to self report. All participants except two members of the older group were currently using a mobile phone. On average, the older adults had used a mobile phone for 9.5 years ($SD=4.8$ years) compared to 8.5 years ($SD=1.7$ years) for the younger group. Another indicator for mobile phone usage intensity, the average amount of short messages written per month differed greatly between the older ($Mdn=2$) and the younger ($Mdn=40$) group ($U=34.5$, $z=-4.157$, $p<.001$, $r=-.68$). Asked about their previous experience with touchscreens, all participants except one older adult stated to have used such systems before. In both groups ticket machines and ATMs were by far the most often mentioned previous encounters with touchscreens. The use of touchscreen-operated info terminals ranked next for the older group, while younger participants named laptops, mobile phones and mobile gaming next often. The most striking difference was that the older users had comparable (and even slightly more) previous experience with stationary touchscreen terminals (33 vs. 29 quotations), but fell behind particularly in the field of mobile touchscreen experience (2 vs. 13 quotations). Similar results were obtained for the question pertaining to previous experience with gesture input. Only one older participant indicated to have operated a system by gesture before (a navigation system), as opposed to 6 younger users (32%), but this difference was not significant ($\chi^2(1)=4.08$, $p_{\text{Fischer}}=.090$). Even though both groups had different prior exposure to technology, their general attitude towards it did not differ. None of the four subscales (enthusiasm, competence, positive and negative effects of technology use) of the TA-EG inventory on technology affinity (Karrer, et. al, 2009) showed a significant group difference ($p>.05$). Neither did the questionnaire data on action strategy (FESE-R, Nachtwei 2008b), which probes for biases within the speed-accuracy tradeoff spectrum ($F(1,36)=2.154$, $p=.151$, $\eta^2=.056$). Both groups had similar values for their accuracy and speed orientation, with a slight bias towards accuracy.

5.8.2 One-finger Gesture Performance

For the one-finger gestures, a subset of the gesture set from the previous study was used. As in study 1A, the older users group traced the gesture patterns significantly slower ($M=6.91$ cm/s, $SD=2.33$ cm/s) than the younger comparison group ($M=9.16$ cm/s, $SD=2.23$ cm/s), $F(1,35)=9.022$, $p=.005$, $\eta^2=.205$, and thus **H1** was supported (cf. **Table 5-3**). The average velocity for both groups lies in between that of the small

($v_{old}=5.65$ cm/s, $v_{young}=7.24$ cm/s) and medium screen size conditions ($v_{old}=10.25$ cm/s, $v_{young}=12.61$ cm/s) from study 1A. Even though the gesturing space on the iPod Touch corresponds roughly to the medium sized gesture space, the speed advantages in study 1A for this condition are assumed to be caused by: 1) a slightly larger screen space (3.6 in vs. 3.5 in), 2) the fact that no physical boundaries were given tempted participants to speed up (and resulted in considerably higher amount of boundary violations), and 3) a device that was always stable and operable by index finger, instead of the hand and thumb operating modes in study 1B. The operation posture did indeed significantly influence the speed of gesture performance, $F(2,70)=97.949$, $p<.001$, $\eta^2=.737$. Contrast analysis revealed that while the table and hand posture conditions did not differ, the thumb operation mode caused significantly slower performances than both the table ($F(1,35)=116.22$, $p<.001$, $\eta^2=.769$) and hand ($F(1,35)=138.09$, $p<.001$, $\eta^2=.798$) conditions. Hypothesis **H7a** (one handed gesture performance is slower) is thus substantiated.

Table 5-3. Comparison of dependent variables according to *age group*.
Asterisks mark significant main effects (** $p<.01$, * $p<.05$)

dependent variable:		younger		older		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
error rate	(%)	1.01	1.19	2.37	1.92	*
average velocity	(cm/s)	9.16	2.23	6.91	2.32	**
angular deviation	(°deg)	5.95	1.01	6.38	1.48	
eccentricity		0.59	0.04	0.58	0.06	
line deviation	(mm)	0.39	0.10	0.41	0.09	
ellipse deviation	(mm)	0.51	0.09	0.59	0.13	

Furthermore, there was a main effect of gesture difficulty class on execution speed ($F(1.6, 56.02)=9.022$, $p=.005$, $\eta^2=.205$). However, it was not the class with the gestures that were rated most difficult to be the slowest in general, but the class rated as medium difficulty. An inspection of the individual gestures showed that gesture velocity varies greatly between the individual patterns, and the medium difficulty class happened to be populated by many gestures that included a right angle (cf. **Figure 5-8**), which were at the lower end of the velocity spectrum. Among the fastest gestures were those that consisted of diagonal swipes, the *cross* gestures and the full circle gestures. We did not observe any interaction of age group and difficulty, or age group and posture, indicating that the older group is not likely to suffer more from more difficult gestures patterns or from difficult postures, such as the single handed usage. Hence, with respect to gesture performance speed, no evidence for **H8** could be obtained.

In contrast to study 1A, a main effect of age group on performance errors could be measured, $F(1,34)=6.712$, $p=.014$, $\eta^2=.165$. Errors comprised those gestures that were either completely missing, had too few sample points, had missing (or surplus) segments or where the fitting algorithm failed. On average, older adults performed an erroneous gesture more than twice as often ($M=2.37\%$) as compared to younger adults ($M=1.01\%$). Remarkably, these values are for both groups below the average error rate of the medium screen size in study 1A, indicating that the operation posture manipulation did not create additional difficulty with respect to error rate. There was also no interaction between age group and posture; therefore, errors were induced in the different posture conditions to a similar degree for both age groups. As expected, there was also a main effect of operation posture ($F(1.57, 53.59)=7.509$, $p=.003$, $\eta^2=.181$) on the error rate, with the highest error rate evoked by the thumb performance ($M=2.73\%$), and the smallest observed in the handheld condition ($M=0.83\%$). Error analysis was conducted only over 37 participants, since one participant of the older group was excluded after outlier inspection. This participant showed an error rate of 19.3% (most of the errors being completely missed gestures), which was close to four standard deviations above the mean in the older group, while all others showed error rates well below 10% (7.04% being second highest).

With respect to the accuracy of gesture execution, results were similar to the first study. For the measures of *form stability* (angular deviation, eccentricity), no main effect of age group, nor interactions of difficulty or posture with age group were observed. Thus, no indications were obtained that older adults had larger problems of fine motor control to accurately reproduce the gesture shapes than their younger counterparts (**H2**). Also for the *direction stability* parameters, no systematic age effect could be found. If the ANOVA is taken to be robust against a slight deviation from normal distribution in the younger group ellipse deviation data ($D(19)=.199$, $p=.047$), the older group can be believed to have greater difficulties to smoothly trace a circular pattern compared with the younger group, $F(1,35)=4.786$, $p=.035$, $\eta^2=.120$. However, non-parametric analysis reveals only a marginable difference ($U=107.00$, $z=-3.48$, $p=.053$, $r=-.32$). For linear gestures, no main effect of age group or interactions with age could be observed. Taken together, across all accuracy measures, no indications of a systematic and substantial drop in accuracy of gesture execution could be put on record, which aligns with the results from study 1A and previous reports (cf. discussion on speed-accuracy tradeoff in Section 5.6.2). Thus, no strong indication in support of **H2** was obtained. Furthermore, even though all accuracy measures were influenced by operation posture (in support of **H7**) and the difficulty level, no interactions with age group were found. Therefore, the hypotheses that older adults would perform more complex gestures less accurately (**H3b**), or that they would have greater difficulties than younger users in performing gestures with their thumb (**H8**), could not be substantiated with the present data.

5.8.3 Two-finger Gesture Performance

The two finger gestures were analyzed separately, since they consisted of a different gesture set and no manipulation of difficulty level was performed for this group of gestures. Since single-handed operation was not feasible for the two-finger gestures, only a 2 (AGE GROUP, between) $\times$ 2 (POSTURE, within) mixed ANOVA was conducted. Descriptive results for the two age groups and the two posture conditions are summarized in **Tables 5-4** and **5-5**.

Table 5-4. Comparison of dependent variables according to *age group*.
Asterisks mark significant main effects (** $p < .01$, * $p < .05$)

dependent variable:		younger		older		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
error rate	(%)	8.11	5.95	14.91	10.93	**
average velocity	(cm/s)	6.25	1.80	5.97	2.73	
angular deviation	(°deg)	4.21	0.82	4.85	1.23	
eccentricity		0.75	0.06	0.73	0.06	
line deviation	(mm)	0.48	0.09	0.59	0.16	*
ellipse deviation	(mm)	0.35	0.06	0.45	0.11	**

Table 5-5. Comparison of dependent variables according to *operation posture*.
Asterisks mark significant main effects (** $p < .01$, * $p < .05$)

dependent variable:		table		hand		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
error rate	(%)	9.21	10.17	13.81	9.92	**
average velocity	(cm/s)	6.97	3.65	5.26	2.23	*
angular deviation	(°deg)	3.93	1.21	5.14	1.47	**
eccentricity		0.72	0.08	0.77	0.07	**
line deviation	(mm)	0.43	0.21	0.64	0.15	**
ellipse deviation	(mm)	0.38	0.10	0.41	0.13	

In general, performance differences according to age groups were not strongly pronounced even for the two-finger gestures. In contrast to the one-finger gestures, no significant reduction in gesture execution speed for the old age group was observed ($F(1,36)=.139$, $p=.711$, $\eta^2=.004$, even though older adults were again slower ($M=5.9$ cm/s) than the younger group ($M=6.5$ cm/s). As for the single-finger gestures in this study, the older group performed significantly more erroneous gestures (14.9%) compared to the younger group (8.11%), $F(1,36)=5.670$, $p=.023$, $\eta^2=.136$, but this effect was not influenced by different operation postures. Noteworthy is also the drastic increase in error rate (by 12.1% for the old group, and 7.1% for the young group

respectively) from single-finger to two-finger gestures, meaning that an erroneous two-finger gesture occurred 6-8 times more often than an erroneous one-finger gesture. This can be accounted for by the increased number of error sources: sometimes only one of the finger traces was present, sometimes both traces were performed sequentially with one finger rather than simultaneously with two fingers, the chance of wrong direction doubled, as did the chance of a failed ellipse fitting (e.g. for the two-finger rotation gestures). A closer inspection of the different error sources shows that the older group had particularly more cases where a gesture was missing completely and even more problems to perform a gesture correctly using both fingers, while circular gestures were less problematic in the group as compared to the younger group (Figure 5-10).

Figure 5-10. Error rates (percentage of total gestures) according to different error sources for both age groups

Moreover, the frequency of omission of the second finger almost doubled for the older group from table (4.7%) to the hand condition (8.5%), while it remained relatively stable for the younger group (1.5% and 2.3% respectively). Vice versa, for the ellipse fitting errors, the error rate doubled for the younger group (table: 3.6%, hand: 7.1%) while it stayed comparable in the older group (3.4% and 3.1% respectively).

With respect to gesture performance accuracy, results differed between *form stability* and *direction stability* measures. The variables indicating how well the basic form was preserved – angular deviation and eccentricity – showed no difference between the age groups, whereas the measures of path smoothness – line deviation and ellipse deviation – both revealed a main effect of age group, $F_{\text{line}}(1,36)=6.091$, $p=.018$, $\eta^2=.145$, $F_{\text{ellipse}}(1,36)=11.486$, $p=.002$, $\eta^2=.242$. Since this was not observed for single finger gestures, it might well be that the coordination of two fingers to perform straight and smooth trajectories caused more problems to the older group than to the younger group.

Apart from the differences between the two age groups, the manipulation of device posture significantly influenced the way the gestures were performed, showing that holding the device in the hand (as opposed to having it fixed at the table) resulted in slower ($F(1,36)=7.192$, $p=.011$, $\eta^2=.167$), less accurate (e.g. for angular deviation: $F(1,36)=21.570$, $p<.001$, $\eta^2=.375$) and more error-prone ($F(1,36)=14.260$, $p=.001$, $\eta^2=.284$) gestures (cf. **Table 5-5**). This clear differentiation between the table and hand operation posture conditions had not been observed for the single-finger gestures. The reason behind this difference could potentially be found in the difference between bimanual and unimanual coordination. For single-finger gestures, each gesture could be traced with the index finger of the preferred hand, irrespective of whether the device was mounted on the table or held in the other hand. In the case of the two-finger gestures, the patterns could be drawn with both index fingers in the table condition, whereas two different fingers of the same hand had to be used in the hand condition, reducing the efficiency and effectiveness of gesture performance.

5.8.4 Subjective Gesture Evaluation

After each gesture tracing part, participants ranked the gestures according to their perceived difficulty. Both groups showed consensus in their rank ordering ($r_s=.902$, $p<.001$). The ranking of perceived difficulty of single-finger gestures showed furthermore great consistency with the ranks assigned in study 1A. Again, simple linear gestures were perceived as easiest, with a tendency to prefer up-down and right-left over the opposite directions. On the lower end of the spectrum were again gestures with many segments, and the unfamiliar directions of the *cross* and *plus* gestures with unfamiliar directions (cf. **Figure 5-11**).

The ranking also supported the effectiveness of the difficulty classes introduced in study 1B: Seven out of the ten easiest rated gestures stem from the easiest difficulty class, and nine out of the ten most difficult rated gestures came indeed from the difficult gesture class. The largest difference in mean rank between the age groups was observed for the *check mark* (✓) gesture. This gesture was ranked significantly higher in the younger group than in the older group.

Figure 5-11. Mean-centered gesture ranking according to the averaged perceived difficulty.
Top: Ten best ranked gestures; *Bottom:* Ten worst ranked gestures

For the two-finger gestures, both groups judged the parallel swipe gestures as easiest, and again a preference of top-down and left-right over the opposite directions was observed (**Figure 5-12**). Most difficult were swipes where the fingers had to be moved in opposing directions, and rotation gestures where two fingers rotate around each other. The highest difference between age groups was observed for the *pinch* and *spread* gestures. This effect can almost certainly be attributed to the higher level of familiarity of the younger group with these gestures that have been advocated by Apple and have become the de-facto standard for zooming in and out on mobile multi-touch devices.

Figure 5-12. Two-finger gestures ranked according to the averaged perceived difficulty

Since the single-finger and two-finger gesture sets consisted of different patterns, the comparison of the behavioral data on speed and accuracy is tricky and makes it difficult to quantify effects. On a descriptive level, the single-finger gestures were performed faster, with fewer errors, and with greater directional stability. On the contrary, form stability was superior in the two-finger condition, yet the patterns in the set did not comprise any gestures made up of several segments, and no acute angles had to be recreated. When asked to evaluate the mental and physical demand of the single- and two-finger gesture tracing task with the help of the NASA-TLX scales, both the mental ($F(1,36)=16.331$, $p<.001$, $\eta^2=.312$) and physical ($F(1,36)=12.744$, $p=.001$, $\eta^2=.261$) demand were rated higher in the two-finger condition. In both cases, mental demand was lower than physical demand. Most interestingly, both age groups rated the perceived demands similarly, since no main effect of age group or interaction of task type with age group was found.

After each of the two drawing tasks, participants were asked to indicate their preferred operation posture of the device. The unimanual thumb-based operation mode was preferred least of all by both groups (**Figure 5-13**). While the younger users preferred to hold the device in their left hand rather than have it fixed on the table ($\chi^2(2)=10.84$, $p=.004$), the elderly users were indifferent in this respect ($\chi^2(2)=5.20$, $p=.093$) as far as the single-finger gestures were concerned. However, both groups preferred the table condition over the handheld operation posture when they had to perform two-finger gestures, ($\chi^2_{\text{old}}(1)=4.77$, $p=.049$, $\chi^2_{\text{young}}(1)=8.0$, $p=.008$).

Figure 5-13. *Left*: Preferred operation posture for single-finger gestures;
Right: Preferred operation posture for two-finger gestures

5.9 Discussion Study 1B

5.9.1 The Influence of Posture and Difficulty on Gesture Performance

The way the device was held showed a clear influence on how well gestures could be executed. The thumb-based operation has been identified to be the most detrimental to faultless gesture input, in support of **H7**. Not only were gestures performed markedly slower by the thumb (velocity reduced by 1/3 in comparison to the handheld and table conditions, **H7a**), but also the most errors were performed in this operation mode, and gestures were performed least accurately on all accuracy parameters (**H7b**). Given the known restriction of thumb flexibility, rotation angle, and the difficulty to reach certain screen regions with the thumb (Hirota, 2003; Karlson, et al., 2006b), this result is not surprising. For example, the difficulties to accurately perform a straight NW $\leftrightarrow$ SE movement with the thumb are clearly visible compared to the execution by index finger (**Figure 5-14**).

Figure 5-14. Finger movement trajectories collapsed over all participants (older group) for each operation posture condition

On the background of a general desire for one-handed interaction (Karlson, et al., 2006a, 2006b), it is nevertheless noteworthy that the one-handed, thumb-based interaction style was clearly the least preferred, not only in the older group, but also among the younger participants, who are sometimes referred to as “Thumb Generation” due to their proficiency in typing on mobile devices with their thumb (e.g. M. Brown, 2008). It seems to be the case that even for relatively simple gesture patterns, like the ones tested in this study, it is not self-evident that they can be performed to the users’ satisfaction in circumstances where only one hand is available.

For the one finger gestures, no difference in errors, speed or accuracy was detected between the table and the handheld posture conditions. This result is understandable if one considers the similarity of how the device is operated in these conditions. In both cases, participants were free to draw the pattern with the index finger of their preferred hand. A different picture emerges for the two-finger gestures. Here, two different ways of handling the device were observed: gesture patterns were usually drawn with the index fingers from both hands in the table condition, while two fingers from the same hand were used in the handheld condition. This difference was documented by a 50% increase in error rate in the handheld condition, a significantly slower gesture execution, and a less accurate shape reproduction. While it is not too astonishing to observe better accuracy and fewer errors if the device is stabilized on the table, it was yet surprising to see that the table condition was also the preferred posture for both groups. Even the younger participants, who indicated a strong preference for handheld over table operation for the single-finger gestures reversed their preference for the two-finger gestures. Thus, also the younger participants do not necessarily perceive multi-touch interaction in a mobile context as most natural and intuitive.

Of greater importance for the objective of this investigation was the fact that no interaction of device posture with age group could be measured, neither for single-finger, nor for two-finger gestures. Problems arising from difficult posture thus seem to have similar effects in both age groups. Any difficulties reported in literature for older adults regarding bimanual coordination (Greene & Williams, 1996; Swinnen, 1998) could not be shown to be detrimental for this gesture input task, thus **H4** was not substantiated. On the contrary, if older adults were given the chance to use both hands (table condition), they were, just like the younger group, faster, more accurate and less error prone. Also **H8** – older adults have greater difficulties to perform gestures in one-handed operation mode – could not be confirmed.

The division of the gesture patterns into three difficulty classes proved to be effective, since a main effect of difficulty could be demonstrated for speed and all accuracy measured. The classes, based on the subjective user rating of perceived difficulty from study 1A, therefore appear to be more coherent. They grasped the true difficulty better than the simple complexity manipulation from study 1A, which did not result in consistent effects. The repeated subjective ranking furthermore showed good consistency with the previous ranking, since 73% of the gestures were ranked in the same class as previously. Moreover, a large between group consistency ($r=.90$) proved that there is a common understanding of what makes an easy and what a difficult gesture. However, contrast analysis revealed that the class with the presumably most difficult gestures did not always lead to the least accurate results, and supports evidence from study 1A that individual gestures show large differences in

their tendency of accurate reproduction, and the difficulty effects are largely dependent on the specific gesture class composition.

5.9.2 The Influence of Age Group on Gesture Performance

The results brought evidence that, as was demonstrated already in study 1A, older adults are slower in performing gestures than the younger comparison group, in support of **H1**. The slowing factor for the single-finger gestures was almost identical to the one identified in study 1A (1.32 vs. 1.30), and comparable, albeit at the lower end, to other reports in the literature (cf. discussion in Section 5.6.2). Velocity differences were still significant, but a little less pronounced for the two-finger gestures (slowing factor 1.05). Comparing the single-finger and two-finger velocities, the younger group sacrificed more of their speed advantage (-32%) than the older group (-14%). On this basis, there is again little reason to assume that older adults have greater difficulties with more complex gestures, as was postulated in **H3**.

However, the older group did produce significantly more erroneous gestures than the young group, for both single- and two-finger gestures. In study 1A, there had also been an almost doubled error rate for the older participants, which had not reached a significant difference at $\alpha=.05$, due to huge intra-group variances. Similarly, in study 1B, the older adults' error rate was about twice as high as the younger ones' across all gestures. It has already been discussed that the sources of errors also differed between age groups, and the older participants' problems were mainly found in a higher rate of totally missed gestures, and a higher rate of missing fingers. From these observations, it seems rather unlikely that the erroneous gestures were produced due to a higher noise in their motor system, problems with coordination, force regulation or endpoint control (Ketcham & Stelmach, 2004). Rather, difficulties might have occurred on a conceptual or strategic level. For example, a common observation was that older participants, unsure whether their pressing of the *proceed* button was effective, pressed again while in the meantime the next gesture had already loaded, thus skipping this trial and missing the gesture. Similarly, instructions to retrace two differently colored lines with different fingers seemed to be temporarily forgotten, which also hints at problems related to memory, attention span and fatigue, rather than insufficient motor performance. This interpretation is in line with the limited evidence for an age-related decrease in gesture accuracy. In the case of single-finger gestures, like in study one, no data in support of **H2** was obtained. Therefore, within the given limitations of the sample and gesture set, there is little reason to assume that age related deficits in motor control lead to a less accurate gesture performance.

For the two-finger gestures, there was a significant difference between age groups for direction stability measures, but not for form stability measures. From a practical point of view, problems with directional stability are of minor relevance compared to problems with form stability: if the basic pattern shape is preserved, a pattern matching gesture detection algorithm could still function sufficiently well, even if the trajectories are not straight and smooth, but contain slight curvatures or jittering. Nevertheless, the greater difficulties observed for older adults to produce smooth two-finger gestures could – at least partly – be due to their increased difficulty to coordinate thumb and index fingers, as reported by Spirduso and Choi (1993). Since the exact choice of fingers was free in this study and not controlled for, no definite answers in this respect can be given.

Taken together the results of single- and two-finger gestures, no systematic or substantial age-related decreases in accuracy were observed. For two-finger gestures, the measurable deviations were still on the scale of less than 1mm, which seems within tolerable boundaries of real applications.

Most interestingly, no interactions of gesture difficulty or device posture with age group could be determined in both tasks for any of the dependent variables. This means that even as gestures became more difficult, or as the posture changed from a comfortable table operation to a more complicated unimanual operation, older and younger participants were not affected differentially, but suffered comparably from these manipulations. Hypotheses **H3**, **H4**, and **H 8** could therefore not be substantiated based on the present results. Therefore, no restrictions on gesture difficulty level, or restrictions on difficulty level with regard to a certain operation posture need to be issued for the older generation. The gist of this study, as was already documented in study 1A, is that older adults perform the gestures slower, but any potential limitations of their motor performance do not seem to hinder them from retracing the patterns as accurately as the younger group. Furthermore, they do not suffer differentially from more complex gestures or device operation postures. Only the two-finger gestures still evoked more problems in the older generation, but this might be compensated with adequate training.

5.10 Conclusions: Age effects on Gesture Execution Performance

This chapter centered around the question whether older adults can perform gestures as a novel input paradigm for technology, and for mobile devices in particular, as well as younger people, or, whether well-documented age-related changes in motor control and manual dexterity cause substantial difficulties towards efficient

and effective interaction. Furthermore, we wanted to find out whether certain interaction scenarios are potentially more detrimental than others, and whether it is possible to establish boundary conditions that allow for satisfactory gesture interaction for older users, and separate these from less favorable interaction conditions. To this end, younger and older adults have been compared in two controlled studies on a range of parameters concerning their gesture execution performance, under varying interaction conditions. The size of the gesture space, the operation posture of the gesture-operated device, the pattern complexity as well as the amount of fingers needed to draw the gesture were manipulated. In this first stage, gestures were analyzed purely regarding their syntactic characteristics, completely detached from semantics and from actual context of use.

The general picture emerging from both studies is that older users do perform gestures slower than younger ones, but no particular problems in accurately retracing the given gesture patterns became apparent, which could be traced back to problems in their motor abilities. On the basis of the results presented above, where most null hypotheses could not be rejected, it cannot be stated that older users do *not* differ from younger ones with respect to accurate gesture performance. If the goal had been an investigation of basic research to deepen the understanding of age-related changes in motor control with regard to a specific motor task, conclusions at this point would be very limited. However, the presented studies should be understood from an applied research perspective, and their explorative nature recognized: no studies exist yet that have documented any interaction gesture performance with older adults. Hypotheses were generated on the basis of very general findings and by analogy from similar tasks. On this background, the obtained results are encouraging with respect to older adults' use of gesture input devices.

With all befitted care to draw any conclusions from predominantly non-significant effects, it still seems to be the case that older adults, at least in the age-band of 60-75 studied here, do not experience any profound difficulties in correctly drawing gesture patterns on a touchscreen surface. Most notably, even in a very restricted gesturing space (simulating a really small screen size), even in a difficult operation posture condition where only thumb-based input was possible, and even involving rather complex patterns consisting of several segments or strokes, no substantial performance decline of the older group compared with the younger group was detected. On the one hand, the absence of performance decline under these potentially difficult conditions increases the plausibility that with simple patterns, sufficient screen size and free choice of operation posture, older adults could perform gesture input well. On the other hand, it is not possible to establish any clear boundary conditions, in which gesture input should be avoided (or under which gesture input is recommended) for older adults, and draw inferences for product design that make a dis-

tion between age groups. Nevertheless, some of the observations lend themselves to recommendations for designers and practitioners.

First, the general slowing in gesture execution that was substantiated for the older generation does, at first glance, not seem to have any implications for design. Efficiency is, in many situations of daily life, not the primary focus of an older user, as long as the goals are achieved. Furthermore, it is often not needed for gesture recognition, since detection algorithms mostly work exclusively on spatial information. However, a system might need to rely on user input speed when thresholds have to be set to discern a single *tap* from a *double tap*, a *simple tap* from a press and hold, or if the gain of *momentum scrolling* has to be set. In such situations, the specifications for older users (e.g. in form of a different configuration file or profile setting) can take the slowing into account. In these two studies, a slowing factor of 1.3 was found for single-finger gestures, and even less for two-finger gestures. Second, concerning accuracy, the only measurable difference along the age groups occurred for the two-finger gestures. In addition, the error rate rose close to 15% for this type of gestures (as opposed to 2.4% and 4.5% for the single-finger gesture). Thus, if design decisions have to be taken in favor of or against a multi-finger gesture implementation for a given function, the potential difficulty for an older user should be considered. Recommendations concerning the aptness of individual gestures are difficult, since no quantitative data is available. However, based on user rankings and subjective reports, older users seemed to perceive such gestures easier that have a strong iconicity and semantics, gestures that are meaningful symbols, which was not observed in the younger users. Contrarily, circular patterns seem to be perceived differentially easier by the younger group. More research is needed yet to substantiate these observations.

In this first conceptualization stage (see Section 4.3.2), a match between the physical characteristics of the older users regarding their motor abilities, and the syntactic characteristics of a realistic gesture set was attempted. The requirements of this input style regarding motor control and manual dexterity do not seem to surpass the older users' abilities, which is encouraging in two ways: first, the ground is laid to build on these results and investigate further the semantics of gestures, in order to develop a gesture set which is not only feasible to execute, but also intuitive for older users. Second, beyond the absence of a substantial performance decline, slowing factors were found to be rather low compared to other input devices, and a majority of older adults stated the preference to use gesture technology in the future. It is therefore worth further pursuing the question whether this novel input technique can contribute to achieve more intuitive technology interaction for an elderly generation, and ease some of the still dominating difficulties while operating technology in general.

6 Gesture Semantics: Which Gestures are Intuitive to Older Users?

The first stage of conceptualizing suitable gesture interaction for older users yielded promising results. No major restrictions of gesture syntax with regard to the older adults' motor abilities were required. However, gestures were tested without semantics and context. The second stage of conceptualization (cf. Section 4.3.2) aimed at exploring the design space for gestures which are also *intuitive* for older users, i.e. where the link between a gesture pattern and the desired action is immediately evident, and no serious effort is required to learn this association. To find such gestures, two studies were conducted.

Study 2A was aimed at identifying intuitive gestures by asking the users directly, and letting them propose their own solutions. Letting the users *participate* in the design process opens up the solution space for gestures which might not have come to mind to a designer who might focus on reliable recognition or other technical problems. While study 2A was explorative in nature and conceptualized to open up the design space as wide as possible, study 2B was set up to narrow down the gesture space again, and arrive at a converging set of gestures which have been empirically shown to be the *best guess* of what might work out as an intuitive interaction language for older users. These gestures were subsequently implemented and tested with a functional prototype (Chapter 7).

Before both studies are described in closer detail, the term *intuitive*, its scientific basis and the way it is understood within the scope of this thesis are explained (Section 6.1). Moreover, the participatory approach taken in these studies is further explained and justified (Section 6.2).

6.1 What is meant by „Intuitive“?

6.1.1 Definition of Intuitive Use

The literature on gesture interaction often praises this type of interface as particularly „intuitive“. This term has many connotations: an intuitive interface is supposed to be easy to use, easy to learn, requires little to no previous knowledge, and just feels “natural” in its use.

Two research groups have already attempted to grasp the buzzword *intuitive use* of products and evaluated its value as a scientific term (e.g. Blackler, Popovic, & Mahar, 2002; Mohs, et al., 2006). In their definitions, the *unconscious application of prior knowledge* is considered as the key feature of intuitive interaction. While Blackler and colleagues additionally stress the *fast* nature of intuitive interaction, Mohs and colleagues add the requirement for the interaction to be *effective* (Blackler & Hurtienne, 2007). A central point in these definitions is that intuitiveness of use is not a characteristic inherent in the product specifications, but emerges through interaction and is largely dependent on the user. Therefore, it is also more appropriate to speak of an *intuitive use* of a product, or an *intuitive interaction*, rather than an intuitive product itself. Within the scope of this thesis, when we speak of intuitive gestures, we speak of gestures which can be used intuitively – in the sense specified above – by a maximum number of users.

Both groups further agree that intuitive interaction can best be described as a continuum that stretches along the dimensions of *product novelty* and *cognitive effort* (Blackler, Popovic, & Mahar, 2005) or the *type of prior knowledge* employed (Naumann, et al., 2007). This knowledge continuum assumes innate knowledge as the most basic and prevalent type of knowledge, followed by sensorimotor knowledge, culture specific knowledge and, lastly, expertise in a specific domain. While intuitive interaction can be found on each level of the knowledge continuum, Hurtienne and Blessing (2007) argue that knowledge from the lower levels of the continuum is more likely to be applied subconsciously. Hence, user interfaces that tap into lower knowledge levels are more intuitive to use. In addition, these interfaces would reach a larger number of users, and hence they also would ensure a more inclusive interface. Gestural interfaces bear the potential to make extensive reference to the cultural and sensorimotor levels of prior knowledge, and thus, in theory, could be a means of making interfaces more intuitive and more inclusive at the same time.

This point is exemplified in **Figure 6-1**, using the shuffle function of a portable music player. Whereas the 2nd generation iPod shuffle involved a switch button to shuffle and provided additional symbols used from other media players, the 3rd generation iPod shuffle uses an unlabeled switch, which requires knowledge of this particular product. The latest iPod nano (4th generation) switches into shuffle mode by shaking the device. It taps into the basic sensorimotor knowledge that shaking a container with (previously aligned) objects mixes them up.

Figure 6-1. Comparison of traditional input to gesture-based input for the *shuffle* function of a portable music player (iPod shuffle 2nd & 3rd generation, iPod nano 4th generation). Different implementations require different levels of prior knowledge from the knowledge continuum (adopted and modified from Naumann, et al., 2007).

In the case of finger-gestures on touch sensitive surface, intuitive gestures are harder to identify, since prior knowledge of such finger movement patterns is hardly generated in other contexts. Still, it can be assumed that indexical and manipulative gestures make reference to previously obtained sensorimotor knowledge, and iconic as well as symbolic gestures ideally tap into culture specific knowledge to achieve a high degree of intuitive use.

6.1.2 Intuition vs. Prior Experience

Prior experience has been shown to be an important predictor of successful product use (Freudenthal, 1998; Langdon, Lewis, & Clarkson, 2007). However, this concept has not yet been defined clearly. In the context of product use, it is usually equated with technology familiarity and assessed through questionnaires, e.g. based on the frequency of exposure and branding familiarity (Langdon, et al., 2007), or level of exposure and depth of knowledge (Blackler, et al., 2002). Linking the concepts of prior experience and intuitive use, the crucial point seems to be how much prior knowledge is generated through prior experience, and whether this knowledge can be activated subconsciously in order to allow for an intuitive use of products. Thus, while prior experience with similar products is surely a common constituent of intuitive use, it is neither a necessary prerequisite, nor is intuitive use the mandatory result of it: First, prior experience as measured by technology familiarity is usually contextualized with the specific product or product class. Intuitive use of products,

however, can also result from a successful application of prior knowledge of cultural, sensorimotor or innate levels (Naumann, et al., 2007). Second, especially in complex scenarios or changing contexts, even high levels of prior experience might not generate knowledge that can be tapped into subconsciously and effortlessly.

6.2 Participatory Design: Involving the User

The notion of *Participatory Design* (PD) refers to a special approach towards designing technical systems, in which the people destined to use the system play a crucial role in the design process. Since the first steps towards including the user in the design process of information systems in the 1970, their role has undergone a gradual change. From viewing the users as victims in need of support in the 1970s, to competent practitioners in the 1980s, to serious professionals in the 1990s, and to valuable sources of information today (Kanstrup & Christiansen, 2006). The PD approach assumes that users are in the best position to determine how to improve their work life and their system interaction experiences, since they are the ones with the most knowledge about what they do and what they need; designers have merely the role of technical consultants (Schuler & Namioka, 1993). Bergvall-Kåreborn and Ståhlbrost (2008) identified three main motives behind the PD approach: First, an *ethical*, democratic perspective, which gives the users the right to influence technological decisions affecting their private and professional life. Second, a *curiosity* perspective, in which one assumes benefits from explorative fathoming of the design space, from knowledge sharing and from mutual learning between the user and designer. Thirdly, an *economic* or *pragmatic* motive, where, in its core, the belief is shared that the actual quality of a product as well as its cost-effectiveness get improved when the designer has access to a users' knowledge, skills and expertise.

Leaning on these motives, the group of older users lends itself particularly for a participative design approach, since this user group is still often disregarded as potential users of technology (*ethical*, cf. Section 2.5), since the ways they use technology and the conditions under which they are willing to adopt technologies are still scientifically debated (*curiosity*, cf. Section 2.4), and since manufacturers begin to realize that there is profit behind designing easy-to-use technical systems, which are attractive not only to the elderly (*pragmatic*, cf. Section 2.6). In addition, the domain of mobile technology and gesture interaction itself can benefit from a participatory design approach. In particular, the relatively new paradigm of interfacing a technical system by means of gestures opens up a vast design space, where commonly accepted interaction patterns are only beginning to be established (Saffer, 2008). Involving the user in the design of interaction gestures seems highly advisable to fill the blank

page, and to arrive at products whose interaction is truly intuitive and which have a chance to be accepted by the end users.

Indeed, much research is currently carried out that involves the users to identify a suitable set of gestures for various tasks and applications. Three-dimensional, user-derived gesture sets have been proposed, for example, for tabletop-mobile phone interaction (Kray, Nesbitt, Dawson, & Rohs, 2010), deformable displays (S.-S. Lee, et al., 2010), auditory interfaces (Wolf, Dicke, & Grasset, 2011) and mobile interaction (Ruiz, Li, & Lank, 2011). Two-dimensional surface gesture sets have been derived from user suggestions for gesture-based text input (Wobbrock, Aung, Rothrock, & Myers, 2005), tabletop applications (Wobbrock, Morris, & Wilson, 2009), culture specific mobile interaction (Mauney, Howarth, Wirtanen, & Capra, 2010; Thalmeier & Koller, 2009), car applications (Döring, et al., 2011), and even interfaces for blind people (Kane, Wobbrock, & Ladner, 2011). Common to this research is the basic idea to let users express their own suggestions for appropriate gestures, which are subsequently categorized, quantified, and re-evaluated. Study 2A and 2B follow this line of research with the focus on intuitive gestures for mobile interaction, and a particular emphasis on appropriate gestures for older adults.

6.3 Study 2A: User-defined Gestures for Mobile Interaction

The first study in this stage (study 2A) was designed around the question whether and how age might influence the kind of gestures that users would preferably perform to operate a mobile device. Identifying differentiating dimensions would enable the development of guidelines for the design of a multi-touch interfaces especially for or including older users. The research questions guiding this study were in particular:

1. Which gestures do younger and older users spontaneously generate to perform a range of typical tasks on a mobile device?
2. What are differences and similarities in the proposed gestures between the two age groups?
3. Which gesture characteristics prevail especially within the older user group?
4. Is it possible to identify gesture patterns that show considerable agreement between participants so they can be used to derive design recommendations?
5. What is the influence of prior knowledge and technology experience on the kind of gestures that users propose, and how do these relate to age?

To this end, older and younger users were prompted to perform gestures with their fingers on a paper mockup mobile phone, which they conceived most suitable for a

range of typical tasks. Since participants were instructed not to ponder extensively over an appropriate solution, but just to perform any gesture which immediately springs to mind, we strived to identify patterns which can be considered intuitive to them. The gesture patterns were subsequently categorized and quantified for the respective age groups.

6.4 Methods Study 2A

6.4.1 Participants

Central focus of this study was the comparison of two age groups. The average age within the older group was 66.4 years ($SD=3.4$), whereas the younger age group was 24.9 years on average ($SD=3.2$ years). The total number of 53 participants included 26 older adults (13m, 13f) and 27 younger adults (14m, 13f). None of them had participated in study 1A or 1B before. While all of the younger participants were using a mobile phone, the proportion among the elderly was only 85%. All but one participant were right-handed; all participants had normal or corrected-to-normal vision, and reported no impairments of wrist or finger flexibility. Participants received a small financial reward.

6.4.2 Apparatus

The study was conducted using a custom-built handheld paper mockup (127 x 77 mm “screen size”) with a USB webcam (Logitech Quickcam deluxe for Notebooks, 1.3 MP) attached to an aluminum arm protruding from below the base element (**Figure 6-2, left**). A USB powered LED light was mounted to the arm to provide lateral illumination of the surface in order to reduce reflections. Due to its lightweight construction, the mockup could be effortlessly held in one hand and freely moved around. The base itself held a stack of 42 paper cards depicting the tasks (see Section 6.4.4) and was covered by a removable Plexiglas insert. Each task card depicted a simplified screen onto which the participants had to perform an appropriate gesture, which was recorded by the webcam. Besides the handheld mockup, the test setup included also a 17” TFT monitor, mounted on a table in front of the participants, on which the current task instructions were displayed. Task instructions were depicted in form of *before* (left) and *after* (right) screens (**Figure 6-2, right**), and accompanied by verbal instructions of the experimenter. The *before* screen always matched the screen depiction of the current task card in the mockup.

Figure 6-2. *Left*: Handheld mockup; *Right*: Example instruction as seen on the monitor, depicting before (left) and after (right) screen.

The design of the screens followed to a large degree the depictions used by Thalmeier and Koller (2009). The final design of the screens was conceived as a compromise between two conflicting goals: On the one hand, as little graphical detail as possible should be given to avoid any associations or dissociations with the screen designs of any existing products, and to ensure maximum generalizability of the participants' gestures. On the other hand, enough detail had to be provided to portray realistic interaction scenarios and to ensure validity of the produced gestures. **Figure 6-3** shows examples of the task card design, including exemplary gesture patterns.

Task: *select multiple*

Gesture:

Swipe diagonal down across multiple objects

Task: *confirm*

Gesture:

Check mark on object

Task: *zoom in*

Gesture:

Magnifying glass

Figure 6-3. Illustration of screen design and mockup with exemplary gestures.

6.4.3 Procedure

Sessions started with a demonstration of the general capabilities of multi-touch interaction by showing some iPod Touch applications to the participants and letting them explore. Care was taken not to anticipate any of the tasks that were investigated in the following. Next, the experimenter explained the handling of the mockup and demonstrated an example task (setting a clock, see **Figure 6-2**), including video recordings of four different gesture suggestions how to solve this task. These comprised a direct manipulation gesture, an iconic gesture, a symbolic, and a two-finger manipulative gesture, in order to raise the participants' awareness of the possible solution space.

In the following, each task was presented visually to the participants by means of a *before* and *after* screen on the centrally mounted monitor, and each task was explained verbally by the experimenter, avoiding technical terminologies. Participants were instructed to perform a finger gesture on the surface of the mobile device mockup, which, from their perspective, would seem most appropriate to proceed from the current (before) state (corresponding to the image on the task card, **Figure 6-2**), to the goal (after) state. They were asked to tip their fingertips into black paste in order to leave a visual trace of their movement trajectory on the Plexiglas insert covering the card. After each task, the experimenter cleaned the surface and removed the topmost card in order to reveal the next task. Upon completion of each gesture, participants had to indicate their understanding of the scenario and their perceived suitability of their gesture on a 5-point Likert scale. The whole session was recorded with the USB webcam at VGA resolution. Each participant performed 42 tasks. The order of tasks varied between participants to avoid sequence effects.

6.4.4 Tasks

The tasks for which the participants had to produce gestures comprised those generic tasks tested by Wobbrock, Morris and Wilson (2009) which were also applicable to handheld devices, and those investigated by Thalmeier and Koller (2009) which were relevant in a mobile usage context for elderly users, provided they were sufficiently easy to explain. For example, we discarded tasks like *minimize window*, *maximize window* (only really suitable for large screen interactions), *print* or *save file* (less relevant for mobile interactions), and *continuous scroll*, *stop continuous scroll* (difficult to explain with a static mockup). This resulted in a set of 28 generic tasks that were part of one or both of the two previous studies. It was extended by 14 specific application functions which are highly relevant for mobile device interaction, for example in the field of music playback (e.g. *play*, *pause*, *stop* etc.) and telephone functions (e.g. *take*

call, end call, write message etc.). An overview of all 42 tasks that were investigated in study 2A is given in **Table 6-1**.

Table 6-1. Overview of tasks. *Remark:* W) also tested by Wobbrock et al., 2009; T) also tested by Thalmeier & Koller, 2009; A) additional task

function name				remark		function name			remark	
generic functions	selection & control	1	select multiple (grid)	W	generic functions	text editing	22	paste	T	
		2	select single	W			23	cut	T	
		3	select multiple (list)	T			24	undo	W,T	
		4	move object	W,T			25	redo	T	
		5	delete	W,T		basic	26	open menu	W,T	
		6	confirm	W,T			27	home	T	
		7	cancel	W			28	help	W,T	
		8	open	W,T		specific functions	music	29	play	A
		9	close	W,T				30	pause	A
	navigation	10	scroll down	T	31			stop	A	
		11	scroll up	T	32			fast forward	A	
		12	previous page	W,T	33			rewind	A	
		13	next page	W,T	34			next title	A	
	view	14	zoom in	W,T	35			previous title	A	
		15	pan (scroll side)	W	36			volume up	A	
		16	rotate	W,T	37			volume down	A	
		17	zoom out	W,T	phone			38	take call	A
	text editing	18	magnify	W,T				39	end call	A
		19	shrink	T				40	make call	A
		20	select text	A		41	write message	A		
		21	copy	T	42	send message	A			

6.4.5 A Gesture Description Language

In order to identify and quantify gestures from the movement patterns produced by the participants, a formalism was necessary which allowed a precise description of the gesture primitives. Existing gesture description systems were either under- or overspecified for this purpose, creating the need to develop a custom description system. The purpose of establishing this gesture description language was to obtain a formalized yet flexible system of labeling gestures which could then serve as the basis for quantitative analysis.

Saffer (2008) discusses a few common systems of movement notations with respect to their usage for gesture classification, e.g. the *Labanotation* (used in choreography), the *Benesh System* movement notation (used mainly in dance), or the *Eshkol Wachman Movement Notation* (EWMN, mainly focused on limp positions). Common to these description systems is the high complexity of the notation score, requiring extensive training to read and write. Furthermore, these systems are all designed to notate full body movements, and therefore unnecessarily complex for 3D hand gestures, and underspecified for 2D finger gestures. They also do not take into account the interplay with a technical system, e.g. the relation of a gesture to an interface object. Other gesture description formalisms take a purely technical stance, and aim at gesture descriptions that can be parsed and interpreted by a computer. Examples of such gesture description systems are e.g. the Gesture Description Language (GDL) proposed by Echtler, Klinker and Butz (2010), the Common Gesture Language (CGL, Kölsch & Martell, 2006), MURML (Kranstedt, Kopp, & Wachsmuth, 2002) or the sign specification system by Losson and Vannobel (1999). Such systems are generally hierarchical in nature and include a range of detailed metric quantifications (e.g. exact coordinates, motion vectors, rotation angles etc.), which make them inappropriate for a researcher intending to describe gestures on a more conceptual level. In HCI literature, gesture description systems for multi-touch interfaces mostly exist just as a collection of gesture primitives (e.g. Saffer, 2008; Wromblewski, 2010) without further specification of direction, location, object-relation etc., or are described only verbally without a structured formalism (e.g. Mauney & Le Hong, 2010).

A new formalism was established as a tradeoff between the specificity level on which a gesture is described, and the generalizability of the gesture description, in order to subsume similar motion patterns under the same descriptor. It was aimed at providing maximal parsimony, while at the same time maintaining readability, and was strongly influenced by the terminology proposed by Thalmeier and Koller (2009). This *gesture description language* is based on the commonly used gesture primitives (e.g. tap, swipe etc.), but is flexible enough to incorporate also iconic and symbolic gestures. The grammar used to form unique gesture descriptions reads as follows (in Extended Regular Expressions (ERE) notation):

(attr? ((sign char*) | motion)+ dir? finger? (rel_{space} loc)? rel_{temp}? sequence?)+

According to this grammar, a gesture description consists of at least one (+: one or more) sign or motion term (|: either A or B) to describe the central movement pattern, which can be specified in greater detail by a couple of optional parameters (*: zero or more; ?: zero or one), such as the *direction of movement*, the *number of fingers* used to perform it as well as its *spatial relation* to screen objects and *temporal relation* to other gesture parts. The *sequence parameter* allows concatenating multiple description terms if the gesture consists of several independent parts (all elements of the gesture description language are listed in **Appendix Table A.4-1**).

On the most basic level, gestures are described by either a *sign* term, if the movement trajectory on the surface refers to a familiar icon (iconic gesture) or symbol (symbolic gesture), or a *motion* term, if the movement trajectory pointed directly at a screen object (indexical gesture) or signaled a direct manipulation (e.g. rotating the screen through a circular motion). While the set of possible signs is limited only by the user's imagination, the set of motion types is finite and predefined. The motion types employed were precisely defined and in line with existing literature (Saffer, 2008; Thalmeier & Koller, 2009). The only difference to previous descriptions is that we have subsumed all straight movements of a finger on the surface under the term *swipe*, irrespective of the speed of movement or the coupled functionality, and therefore created a unitary term for what has been labeled "slide", "swipe", "flick", "fling", "drag" or "nudge" (e.g. Saffer, 2008). Some examples of gesture names can be seen in **Figures 6-3** and **6-4**.

6.4.6 Gesture Analysis

The 2226 registered motion patterns were manually classified with the help of the recorded screen capture videos and the experimenter's protocol sheets into 568 different gestures. A gesture glossary was created in which each gesture was referenced by a unique *gesture ID*, a *verbal description* according to the description language (Section 6.4.5), and a screenshot of the gesture pattern. Moreover, the gesture glossary contains information on the gesture's basic *type* and quantifies the gesture with the parameters *number of fingers* used, the *number of segments* and *number of strokes* it consists of (**Figure 6-4** shows a screenshot of the gesture glossary).

ID	Gesture	Gesture image	Type	Fingers	Strokes	Segments
41.01	Letter(s) N on object		alphanumeric	1	3	1
41.02	Tap on object		indexical	1	1	1
41.03	Spread horizontal across object		manipulative	2	1	1
41.04	Letter(s) K N on object		alphanumeric	1	1,5	4
41.05	Swipe left from object towards left border		manipulative	1	1	1
41.06	Swipe down across object		manipulative	1	1	1
41.07	Arrow right on object		iconic	1	1,5	2

Figure 6-4. Screenshot showing a section from the gesture glossary. User-derived gestures for the task *write message* are displayed. Each gesture is described with unique ID, verbal description, image, type, number of fingers, number of segments and number of strokes

A *segment* is defined as a continuous finger movement on the screen without intermediate lift-offs of the finger, and can consist of several strokes. A *stroke*, in this context, is a clearly separable finger movement from one resting position of the finger to another. For example, an X gesture consists of two segments made of a single stroke each, while a V gesture consists of one segment, made up from two strokes. The types into which the gestures were classified were derived from the gesture framework proposed in Section 3.2.2. In particular, manipulative gestures were discerned from indexical, iconic and symbolic gestures.

- **Indexical gestures:** The gesture consists of a mere positioning of the finger on a certain screen object or region. A simple tap is the predominant indexical gesture.
- **Manipulative gestures:** Screen objects or regions are manipulated with the help of movement patterns such as repeated tapping, long presses, dragging movements, swift swipe or flick movements, spreading or pinching movements, encircling or boxing screen objects etc.

- **Iconic gestures:** Drawings that result in the depiction of real world objects. Common iconic gestures were arrows, but also crosses, houses, magnifying glass, ear, loudspeaker, houses, scissors and other objects were schematically drawn on the screen.
- **Symbolic gestures:** A sign that was neither an iconic object depiction, nor an alphanumeric character, which counted as an extra subset category. Examples comprise e.g. a minus sign, plus sign, bracket, question mark etc.
- **Alphanumeric gestures:** These gestures consisted of one or more letters or numbers which were drawn with the finger on the screen. In few cases, even whole words were drawn.
- **Combined:** When a gesture consisted of separate parts which stem from different categories, they were counted as combined. In most cases, these gestures contained a manipulative and an iconic or symbolic segment.

The gesture glossary was sectioned according to the investigated tasks. Each movement pattern was cross-checked with the already existing gesture entries for this task. If it could be subsumed under one of the already existing entries, it was referenced by existing ID and name; otherwise a new glossary entry was created.

When naming a new gesture entry according to the gesture description grammar, the principle of parsimony was followed: a simpler description was generally preferred if different descriptions were possible (e.g. “tap on multiple objects sequentially” was preferred over “tap on object, then tap on another object, then tap on another object”). Optional descriptive elements were only applied if they were unambiguously evident from the video or protocol sheets, or made explicit by the participants (e.g. tap somewhere on the screen vs. tap on a specific object), if they were needed to set the gesture apart from similar ones (e.g. curved arrow right vs. (straight) arrow right) or generally needed by the gesture itself (an arrow needs a direction specification) or the task (the task *rotate* needs to specify whether a user-generated circular movement is performed clockwise or counter-clockwise).

6.5 Results Study 2A

6.5.1 Types of Gestures

Across all participants and tasks, manipulative gestures were the predominant gesture type (54.7%). Indexical (19%) and iconic gestures (16.6%) were also often suggested, whereas alphanumeric (5.8%), combination (2.3%) or other symbolic gestures (1.7%) occurred only rarely (**Figure 6-5, left**). Older and younger partici-

pants differed in the type of gestures they performed ($\chi^2(5) = 198.808$, $p < .001$). The younger group showed a significant increase in manipulative gestures as compared to the older group ($\chi^2(1) = 151.531$, $p < .001$), while the older participants produced more iconic gestures ($\chi^2(1) = 100.068$, $p < .001$), or alphanumeric symbols ($\chi^2(1) = 42.581$, $p < .001$). The group comparison revealed no differences between older and younger users for symbolic, indexical or combined gestures at the Bonferroni-corrected $\alpha = .0083$.

Looking still deeper into the kind of gestures proposed by the two participant groups, it can be observed that especially gestures involving taps, swipes, double taps and/or multiple fingers, were performed more often by the younger participants as compared to the older ones (**Figure 6-5, right**). Also press & hold, spread, pinch or triple tap gestures occurred more frequently within the younger group. The largest group differences can be seen in the use of double-taps (14.3% younger, 2.7% older) and multi-finger gestures (12.5% younger, 2.3% older), showing that such patterns are not among the first that older people would think of when operating a device through gestures. Conversely, the most frequent iconic gesture (depicting an arrow), was used in 14.3% of all proposed gestures among the older group, and yet only by 2.1% of the younger group. Also all subsets of the alphanumeric gesture category, such as gestures containing numbers, single letters or whole words, show higher frequencies among the older participants.

Figure 6-5. *Left*: Frequencies of gesture types for each age group;
Right: Frequencies of common gesture elements for each age group

6.5.2 Number of Fingers and Gesture Complexity

All participants were familiarized with the multi-finger interaction style on an iPod Touch device before the experiment, and were shown an example of a multi-finger gesture during the instruction phase. They were free to choose the number of fingers to use for performing the gesture. Nevertheless, the vast majority of proposed gestures (89.4%) were performed with only a single finger. The multi-finger gestures were performed mostly using two fingers (9.9%), and only very rarely three or more

fingers were used (0.7%). The maximum number of fingers used simultaneously was five, which occurred only once in the younger group. The distribution of the used fingers differed between the two age groups ($\chi^2(2) = 92.541, p < .001$), with the younger group relying more on multi-finger input than the older group. In particular, the older group has produced significantly more one-finger gestures (96%) in comparison with the younger group (83.1%, $\chi^2(1) = 92.333, p < .001$), while the younger participants suggested more often two finger gestures (older: 3.7%, younger: 15.9%, $\chi^2(1) = 87.716, p < .001$), or gestures where three or more fingers were used (older: 0.4%, younger: 1.1%, $\chi^2(1) = 3.732, p = .053$).

In the study 1A, gesture complexity had been differentiated merely on the basis of lines that the gesture consists of (see Section 5.4.1). This time, for each gesture, first the number of fingers involved was counted. Then, for each finger, gesture *segments* were quantified, and for each segment, the number of *strokes* was counted (cf. Section 6.4.6). A *complexity score* was then calculated for each gesture as *number of fingers* * *average number of segments per finger* * *average number of strokes per segment*. The complexity score of the observed gestures ranged from 1 (e.g. a simple *tap*) to 16 (multiple fingers perform a multi-segment gesture). On average, older participants ($M = 2.32, SD = 0.80$) suggested more complex gestures than the younger group ($M = 1.78, SD = 0.33; t(32.8) = 3.187, p = .003; r = .49$).

6.5.3 Agreement Score

Of central interest in this research is not only to spot potential age-related differences in the preference of suitable gestures, but also to identify mappings between a gesture pattern and the task it serves which are intuitive for a large proportion of potential users and, hence, worth implementing.

The fine-grained gesture description terminology has led to a large number of different gestures identified for each task. The number of different gestures proposed for a particular task varied between nine (e.g. for *play* or *pan (scroll side)*), and 40 (for the task *paste*), with an average of 22.6. These numbers demonstrate the high diversity of the performed gestures, but disregard the distribution of users over the proposed gestures. In order to assess the general agreement between users, we adopted the measure suggested by Wobbrock, Aung, Rothrock & Myers (2005), who, in the context of user-derived gesture quantification, define an *agreement score* A as:

$$A = \frac{\sum_{r \in R} \sum_{P_i \subseteq P_r} \left(\frac{|P_i|}{|P_r|} \right)^2}{|R|}$$

In this equation, R denotes the set of all tested tasks, P_r is the set of uniquely identified gestures for each particular task, and P_i is the number of identical suggestions for each gesture in P_r . The range of A is $[|P_r|^{-1}, 1]$. A higher value of A represents a higher tendency of the participants to settle on the same gesture. Agreement scores were calculated independently for the younger ($A=0.247$) and the older ($A=0.137$) user group, which showed to differ significantly across the investigated tasks ($z=-4.345$, $p<.001$, $r=-.67$). The younger participants were more coherent in the gestures they proposed for a particular task, while the older users showed a larger diversity within the suggested gestures. The agreement differed also greatly depending on the task. The highest overall agreement score ($A=0.47$) was measured for the task to select a single element (e.g. from a list or grid): close to 50% of all users used a simple *tap* to fulfill this task. Other tasks with relatively high consensus on the appropriate gestures include either very generic and basic tasks, such as *select text*, *scroll side*, *open* or *move object*, but also very specific tasks associated with music playback or calling functions, such as *play*, *pause*, *rewind*, *fast forward*, *take call* or *make call*. The agreement scores are depicted for each task and user group independently in **Figure 6-6**.

Figure 6-6. Agreement scores for all tasks, sorted from highest to lowest overall agreement.

6.5.4 Matching Gestures and Tasks

One aim of this study is to identify gestures that spontaneously come to mind as the most suitable for typical interactive tasks on a mobile device, in order to have an empirical basis for the selection of gestures to be implemented in a future functional prototype. Given the generally low agreement on most tasks, selecting a specific

gesture to match a task, which can be considered intuitive to the *majority* of users, becomes difficult (see discussion in Section 6.6.3). For all tasks where the overall agreement was at least .25 (roughly 25% of identical gestures were produced), the tasks are analyzed in greater detail in the following, and different approaches of solving the tasks in the two age groups are pointed out. The pie charts in **Figure 6-7** display the overall percentages of the produced gestures (gestures with less than 5% averaged occurrence collapsed into “other”). Tasks are ordered according to overall agreement (cf. **Figure 6-6**), in descending order.

Select single object	
<p> ■ Tap on object ■ Circle on object ■ Other (< 5%) </p>	<p>The selection of a single object (from a list) was consistently performed most often by a single <i>tap on the object</i> (old: 57.7%, young: 77.8%). Younger users also performed a <i>double tap</i> for selection (7.4%), older users also circled the object (7.7%) or drew a check mark on top (7.7%).</p>
Rewind (Song)	
<p> ■ Swipe left on object ■ Tap on object ■ Swipe left on neutral space ■ Other (< 5%) </p>	<p>The most common gesture to rewind a song was a swipe left on the song progress bar (=dragging its end further left; old: 38.5%, young: 74.1%). The younger users almost exclusively interacted with the progress bar (tapping further left was the alternative gesture, 22.2%), older users also interacted with the surrounding space, e.g. swiped left on neutral space (15.4%).</p>
Select Text	
<p> ■ Swipe right across multiple objects ■ Curve diagonal down from object to object ■ Other (< 5%) </p>	<p>Both younger (59.3%) and older (53.8%) users preferred to swipe right across several words in order to select a text segment. As the text segment had to be selected across a line break, the only other gesture that achieved some consensus was to move down from the beginning of the first word to the end of the last word, imitating a text selection across lines with pressed mouse button (young: 22.2%, old: 7.7%).</p>
Pause	
<p> ■ Tap on object ■ Tap on neutral space ■ Other (< 5%) </p>	<p>Pausing a music playback was mostly attempted by tapping on the play symbol (▶), which was intended to signal the current playback mode, but was conceived by the participants as an interaction object (young: 77.8%, old: 34.6%). While 19.2% of the older users also tapped somewhere else on the screen to pause, only 7.4% of the younger users thought this appropriate.</p>
Fast Forward (Song)	
<p> ■ Swipe right on object ■ Tap on object ■ Swipe right on neutral space ■ Other (< 5%) </p>	<p>The gesture to fast forward a song was generally the opposite of the rewind gesture, thus a very similar picture emerged: Swiping right on the song progress bar was performed by 74.1% of younger users (34.6% old), tapping further right on the progress bar was performed by 22.4 % of younger users (22.2% old), and swiping right somewhere else was only used by older users (11.5%).</p>

Figure 6-7. Selection of tasks (with agreement score >.25) and the corresponding gesture suggestions. Overall frequencies of gesture suggestions are displayed (%)

From the selection of tasks presented above, it can be observed that indicative gestures (mainly simple taps) or manipulative gestures (mostly swipes) were the predominant gestures. Furthermore, even in these cases of relative high agreement, hardly any gesture was suggested by more than 50% of the participants. In addition, for some gestures, the difference between the age groups is substantial: for example, for the tasks *pause*, *rewind*, or *fast forward*, about $\frac{3}{4}$ of the younger group performed the same gesture, while only about half that many older participants settled on the same gesture.

For all tasks, the most frequently suggested gestures in each age group are displayed in **Table 6-2** and **Table 6-3**. For only 15 of the 42 tasks (35.7%) the most frequently suggested gesture coincides between the younger and the older participant group (**Table 6-2**). For the majority of tasks, different gestures were suggested (**Table 6-3**).

Table 6-2. Most frequently suggested gestures for tasks in which older and younger share the favorite gesture. Sorted by gesture frequency in older group (* = first rank shared by more than one gesture)

		older		younger	
task		most frequent gesture	%	most frequent gesture	%
2	select single	Tap on object	57.7	Tap on object	77.8
20	select text	Swipe right across multiple objects	53.8	Swipe right across multiple objects	59.3
38	take call	Tap on object	50.0	Tap on object	59.3
1	select multiple (grid)	Tap on multiple objects sequentially	42.3	Tap on multiple objects sequentially	14.8
33	rewind	Swipe left on object	38.5	Swipe left on object	74.1
5	delete	Cross on object	34.6	Cross on object	40.7
6	confirm	Tap on object	34.6	Tap on object	37.0
26	open menu	Tap on object	34.6	Tap on object	25.9
30	pause	Tap on object	34.6	Tap on object	77.8
32	fast forward	Swipe right on object	34.6	Swipe right on object	74.1
3	select multiple (list)	Swipe down across multiple objects *	26.9	Swipe down across multiple objects	51.9
4	move object	Curve up from one object to another *	23.1	Curve up from one object to another	44.4
31	stop	Tap on object *	23.1	Tap on object *	29.6
17	zoom out	Pinch diagonal	11.5	Pinch diagonal	14.8
19	shrink	Pinch diagonal *	3.8	Pinch diagonal	22.2

Table 6-3. Most frequently suggested gestures for tasks in which older and younger have different favorites. Sorted by gesture frequency in older group (* = first rank shared)

		older		younger	
task		most frequent gesture	%	most frequent gesture	%
29	play	Tap on object	50.0	Double tap on object	70.4
16	rotate	Curved Arrow clockwise	38.5	Quarter-Circle clockwise	40.7
8	open	Tap on object	30.8	Double tap on object	63.0
9	close	Tap on object	30.8	Double tap on object	44.4
11	scroll up	Swipe up across multiple objects	26.9	Swipe down across multiple objects	44.4
12	previous page	Arrow left	26.9	Swipe right	14.8
13	next page	Arrow right	26.9	Swipe left	22.2

15	pan (scroll side)	Arrow left	26.9	Swipe left	66.7
40	make call	Tap on object	26.9	Double tap on object	74.1
7	cancel	Letter(s) N on object	23.1	Cross on object	25.9
21	copy	Letter(s) K	23.1	Double Tap on object	25.9
4	move object	Curve up from one object to another *	23.1	Curve up from one object to another	44.4
31	stop	Tap on object *	23.1	Tap on object *	29.6
24	undo	Curved arrow counter-clockwise	23.1	Arrow left	11.1
19	shrink	Pinch diagonal *	3.8	Pinch diagonal	22.2
10	scroll down	Swipe down across multiple objects	19.2	Swipe up across multiple objects	48.1
22	paste	Swipe right below object	19.2	Double Tap below object	7.4
28	help	Tap on object	19.2	Question mark across multiple obj.	29.6
34	next title	Swipe right on object *	19.2	Swipe left on neutral space *	14.8
35	previous title	Number(s) 2	19.2	Swipe right on neutral space	14.8
41	write message	Tap on object	19.2	Swipe right across object	25.9
25	redo	Curved arrow clockwise	15.4	Swipe right	7.4
39	end call	Tap *	15.4	Cross on neutral space	33.3
42	send message	Tap in object *	15.4	Check mark on object	22.2
14	zoom in	Tap	11.5	Spread diagonal	25.9
18	magnify	Plus sign	11.5	Spread vertical	22.2
23	cut	Scissors	11.5	Cross	11.1
27	home	House	11.5	Cross	14.8
36	volume up	Swipe up from object	11.5	Tap on object, then Swipe up *	29.6
37	volume down	Tap on object	11.5	Tap on object, then Swipe down	33.3

In two tasks, a list had to be manipulated to bring items into the field of view that were either further down or further up the currently visible section. The term *scrolling* or any directional cues (*up*, *down*) were avoided in the verbal instruction. Interestingly, in these situations, participants were far from concordant concerning the direction in which a finger gesture should be performed to achieve the goal. If, in order to fulfill the task, the participant had to *scroll down*, 30.8% of the older group and 55.6% of the younger group suggested a gesture including an upward finger movement, against 50% and 44.4% respectively including a downward finger mo-

tion. Vice versa, for the task of scrolling upwards, 53.8% of the older and 37% of the younger group moved the finger upwards, and 34.6% (older) and 59.3% (younger) produced a downward finger movement. These differences are significant ($\chi^2(2) = 8.699, p = .013$). While in the majority of cases, older participants produced a gesture in the direction they wanted the list to move, the younger users produced gestures in the opposite direction. This observation is further discussed in Section 6.6.1.

6.5.5 Influence of Prior Experience vs. Influence of Age

The above quantifications show that the older group produced quite different gestures compared with the younger group. The reasons, however, might be found to a larger degree in the different prior exposure to and experiences with touch and gesture based technical systems, rather than in any age-related physiological or cognitive changes. Correlation analysis was used to shed more light on this link. Of prime interest in this analysis was to inspect how age relates to a particular gesturing style, and whether any correlation persists after prior experience with technology is accounted for. For each participant, the parameters *computer literacy*, *prior mobile phone experience*, *touchscreen experience* and *gesture experience* were assessed (they are described more detailed in Section 7.3.4). In addition, the gesturing style of the participants was quantified through their average *complexity score*, their average *number of fingers* used, and their relative frequency of producing a certain *gesture type* (e.g. manipulative, iconic gestures etc.). The following sections describe the most interesting correlations. The complete correlation matrixes can be found in **Appendix A.4.3**.

Correlations between gesture parameters. First of all, high correlations between the parameters describing the suggested gestures could be observed. Each gesture type correlated significantly ($p < .05$) with the complexity score: high correlations were observed for manipulative gestures ($r_s = -.72, p < .001$), iconic gestures ($r_s = .71, p < .001$), alphanumeric gestures ($r_s = .61, p < .001$) and combined gestures ($r_s = .55, p < .001$), medium correlation for indexical gestures ($r_s = -.37, p < .01$), and a small correlation for symbolic gestures ($r_s = .28, p < .05$). This means that manipulative and indexical gestures are usually associated with lower levels of complexity, while iconic, alphanumeric, symbolic and combined gestures tend to be more complex. Furthermore, the number of fingers employed is usually higher in manipulative gestures ($r_s = .46, p < .001$), while iconic ($r_s = -.40, p < .01$) and alphanumeric gestures ($r_s = -.23, p < .05$) are negatively correlated because they are usually executed only with a single finger.

Correlations between experience parameters. A similar level of dependency exists between the parameters capturing the participants' prior experience with technology. Besides the correlation between mobile phone experience and touchscreen experience ($p > .05$), all other parameters are significantly positively correlated ($.27 < r < .47$).

Thus, people who score, for example, high on computer literacy, generally also used mobile phones to a higher degree, have had more touchscreen experience and are more familiar with gesture input, and vice versa.

Correlations between age, gesture parameters and experience. Most importantly, the participants' age correlated highly with the types of gestures they performed. There were significant positive interactions of age and the frequency of iconic ($r_s = .46, p < .001$), alphanumeric ($r_s = .40, p < .01$), and combination ($r_s = .27, p < .05$) gestures, and a negative correlation with manipulative gestures ($r_s = -.58, p < .001$), which reflects the results reported in Section 6.5.1. Moreover, also the individual's complexity score correlated with age ($r_s = .48, p < .001$), and the average number of fingers was negatively correlated ($r_s = -.48, p < .001$). In general, age provided the most correlations with the gesture parameters, and also the highest. However, not surprisingly, prior experience with technology correlated strongly negatively with age: computer literacy ($r_s = -.60, p < .001$), mobile phone experience ($r_s = -.70, p < .001$), touchscreen experience ($r_s = -.50, p < .001$), and gesture experience ($r_s = -.42, p < .01$).

Partial correlation analysis. Given the dependencies between the experience and the gesture parameters (e.g. the number of fingers used is positively correlated to all technology experience parameters ($.33 < r_s < .48$), and computer literacy was found to be negatively correlated to the complexity of gestures, $r_s = -.45, p = .001$), partial correlations provided more insight into the relationship of prior experience and age on gesture performance. A first partial correlation analysis determined the influence of prior experience on gesture performance if age is controlled for. While without controlling for age, significant correlations between the experience scores and complexity score, average fingers and the frequency of manipulative gesture existed, only the relationship between gesture experience and number of fingers remained significant in the controlled condition ($r = .35, p < .05$). On the other hand, looking at the influence of age on the types of gestures and control for all four experience parameters, the picture is reversed: Out of six significant correlations in the uncontrolled condition, only the relation between age and number of fingers is no longer significant, all other correlations remain significant ($.31 < |r| < .51$). Taken together, we can conclude that next to prior experience with technology, age is also accountable for the way that gestures are performed. While the number of fingers that participants use for gestures is mostly determined by the degree to which they have used a multi-touch enabled product before, the *type* of gestures they produce, and the degree of *complexity* of the patterns, seem to be determined by age beyond the influence of prior experience with technology.

6.6 Discussion Study 2A

The explorative nature of study 2A complements the hypothesis driven approach taken in the first stage (studies 1A and 1B), and resulted in a wealth of qualitative data on user-generated gestures. A key finding of this study was the observation of differences in gesturing style between older and younger adults, together with the insight that these differences are not only accounted for by different levels of prior experience with technology. Before the scope of the results and their relevance for a prototype design are discussed more closely (Section 6.6.3), two limitations need to be addressed. First, the use of a paper mockup influenced the way how gestures were drawn and limits their generalizability to real devices (Section 6.6.1), and second, the quantitative data is strongly contingent on the chosen classification and analysis method (Section 6.6.2).

6.6.1 Influence of the Paper Mockup

The deliberate choice to present the screen stimuli by means of a paper mockup instead of a real device was taken in order to avoid bias of the participants towards a certain interaction style, and second, to counter expectations that the device would react to the “right” gesture. The lack of immediate feedback, however, might at times have caused misconceptions that would not have arisen in a similar way on a real, reactive device.

For example, the relative high amount of arrows seen in the produced gestures (14.3% of all gestures in the older group), could be a result of the static nature of the mockup. Participants not familiar with the concept of direct interaction (e.g. that a screen object moves directly and instantly with the finger tip on the touch screen surface) could only resort to icons to convey the intended direction of movements. If an object on the screen had started to move right at the beginning of a movement, different gesture patterns would most probably have resulted. It seems particularly likely that older users too would have used patterns of direct manipulation more frequently (at the expense of iconic and symbolic gestures) if the touchscreen had been reactive. This assumption is followed up further in studies 2B and 3.

Similarly, the scrolling direction considered to be appropriate is likely to be affected by the static nature of the mockup. Since the list did not start moving immediately under the finger, any direction of finger movements seemed appropriate and was maintained throughout the study. The underlying issue is that scrolling can be based on two different concepts: the *paper* metaphor, which assumes direct manipulation of a physical list with the fingertip, and therefore the finger – along with the list – has to be pushed *upwards* to see content further down (*scroll down*) and vice versa. This is

commonly used in touch screen devices. Conversely, the *window* metaphor, mainly used on computer screens, assumes a stable list over which the users can pull a moving window (e.g. handled through scrollbars in window-based computer operating systems). Here, the finger has to push (the visible window) *downwards* if content further down should become visible (*scroll down*). Older users have been reported previously to have particular problems with the appropriate scrolling direction (Hanson, 2011). But not only older adults, also younger users, in a similar study, were found to diverge highly for the “scroll down” gesture: half of them moved the finger downwards, while the other half moved upwards (Thalmeier & Koller, 2009).

Mauney & Le Hong (2010) identified ownership of a gesture-enabled device as an explanatory factor, since owners of a finger-driven touchscreen were more likely to adopt the scrolling pattern along the paper metaphor. Our findings suggest that older users most frequently relied on the window metaphor, while younger users predominantly used the paper metaphor. Given the strong negative correlations between age and previous gesture experience ($r_s = -.42$) or touchscreen experience ($r_s = -.50$) in our sample, we can assume that for the older adults it was generally more intuitive to apply knowledge of scrolling behavior from the computer domain, whereas younger users, who were more strongly influenced by prior finger-touch experience, transferred the product specific knowledge intuitively.

Taps were found to be the most frequently performed gesture type (33.1%), closely followed by swipes (32.9%). These results match the findings of Mauney and Le Hong (2010), who also list swipes (32.5%) and taps (22.4%) as the most frequent gestures. The high percentage of simple taps in the present study can be attributed, at least partly, to the same observation that Mauney and Le Hong made: (2010, p. 11) “participants imagined a button or context menu to trigger the required action with a tap, especially when they had difficulty thinking of an appropriate gesture”. This was the case even though participants were instructed to perform gestures only based on the visible screen elements and not imagine other interface objects such as menus or softkeys.

On the positive side, the abstract nature of this static paper mockup also opened up the possibility to imagine gestures beyond some seemingly well-established patterns. For example, the task of *zooming out* received very low agreement scores (cf. **Figure 6-6**) for older *and* for younger users, even though many of them, especially the younger ones, are familiar with the *2-finger pinch* movement advocated by Apple, and currently the quasi-standard for modern touchscreen devices. In our study, in fact only 18.9% of all users used this tow-finger pinch movement (irrespective of direction). By abstracting from a real device, the chosen method elicited a wealth of creative suggestions. Some of them, naturally, lack consensus, but others might be established as

true alternatives to what manufacturers have already proposed, and at the same time match the intuition of older adults better.

6.6.2 Influence of the Analysis Method

The richness and diversity found in the gestures performed must be viewed in the light of the language that was used to describe gestures (Section 6.4.5). Its high degree of granularity accounts for the large number of different gestures identified, and, consequently, for the sparse agreement on a particular gesture for a particular task. Still, the average agreement score measured in this study ($A=0.192$) does not fall far behind the agreement scores observed in related studies. Lee and colleagues (2010), for example, report an average agreement score over all conditions and tasks of $A=0.252$. Thalmeier and Koller (2009) present an (estimated) average agreement score of $A=0.283$, Wobbrock and colleagues (2009) report average agreement scores of $A=0.28$ for two-handed gestures and $A=0.32$ for one-handed gestures. Ruiz, Li and Lank (2011) found an (estimated) average agreement of $A=0.344$ (but only one task above $A=0.45$), and only Kühnel and colleagues (2011) show a slightly higher agreement ($A=0.41$). None of these studies compared between age groups, which makes the particularly low agreement among the older participants ($A_{\text{young}}=0.247$, $A_{\text{old}}=0.137$) an interesting observation, emphasizing the particular difficulty of finding appropriate and commonly accepted gestures for this user group.

6.6.3 Scope of the Results

Two research goals were to be addressed in this study: First, investigating whether there is any difference in intuitive gestures for different age groups, and second, which gestures are most appropriate for certain tasks for older adults. Concerning the first point, differences between age groups were identified on all parameters capturing the general gesturing style: older adults performed different types of gestures, with fewer fingers, produced more complex and more diverse gestures, and these observations are not simply effects of prior technology experience.

Concerning the second point, we obtained a list with a “winning” gesture for each task, based on the frequency of suggestion. Whether these gestures can be regarded as most *intuitive*, or even as most *appropriate*, can be debated though. The study demanded a high degree of ideation from people that were not familiar with the target domain. The benefit of opening up an impartial design space was counteracted by the lack of pragmatic relevance that some solutions exhibited and the poor consensus among participants. Some examples discussed above, e.g. high frequency of arrow gestures, the confusion of scrolling direction, or the expectation of a menu

(Section 6.6.1) are indications that particularly the older participants might have had difficulties in imagining how a real touch and gesture controlled device would function. In this light, some of the proposed gestures (e.g. writing whole words with the finger on the screen) can be considered makeshift solutions rather than the most appropriate. Furthermore, the focus in this investigation was put on gesture semantics, rather than syntax, which led to gestures that might be intuitive (in the sense of an unconscious activation of prior knowledge, e.g. drawing a pair of scissors for the *cut* task), but lack the efficiency that is desirable for tasks that need to be frequently executed. On top of that, agreement on a commonly accepted “best gesture” was extremely low among the older adults.

For these reasons, we sought to validate the results of his study in a follow-up study (2B), before conclusions on the appropriateness of certain gesture types and specific task-gesture matches are drawn for older users.

6.7 Study 2B: Evaluating a User-defined Gesture Set

Study 2B was conceived to validate and refine the results obtained in the previous study. To this end, the most promising gesture patterns from study 2A were re-evaluated by an age-splitted sample using a multiple-choice paradigm. For each task, participants were shown four possible gestures, which they had to rate according to their perceived suitability for this task. One major aim of this study was to separate gesture preferences that were merely owed to the paradigm in study 2A, from gestures that persist as preferred choices for particular age groups. The research questions guiding this study were otherwise very similar to the ones investigated in study 2A:

- Which gestures do older and younger users perceive most suitable for typical mobile interaction tasks?
- Do these gestures differ systematically between the age groups, and if so, what are general differentiating criteria?
- Can the gesture preferences among age groups identified in study 2A be validated in a different paradigm?

The results from this study will then serve as the empirical basis for the gesture set to be implemented and evaluated in a functional prototype (Chapter 7).

6.8 Methods Study 2B

6.8.1 Tasks and Gestures

The set of interaction tasks that were tested in this study comprised 34 typical tasks one might perform on a mobile device, ranging from generic (e.g. *select item*, *scroll up*) to application specific tasks (e.g. *pause*, *send message*). The tested tasks were a subset of the tasks investigated in study 2A (see Section 6.4.4). The tasks which showed less than 5% agreement in the study 2A (average agreement score $A < .05$) were not tested again, since no “best” gestures could be determined for these tasks unambiguously. The tasks discarded that way included *redo*, *undo*, *cut*, *paste*, *home*, *magnify* and *shrink*. *Zoom out* showed an agreement score of $A=.044$, but was left in the set as the pairwise counterpart of *zoom in*; instead, *copy* was also discarded since the semantically connected actions of *paste* and *undo* were no longer part of the set.

For each task, four gestures were selected to be included in the multiple choice test, according to the following selection algorithm:

1. The two most frequently selected gestures of the older group are included
2. The two most frequently selected gestures of the younger group are included
3. If two gestures are equally ranked within one age group, the one with the higher rank in the other age group is selected
4. If the selected gestures are less than four, the next overall frequently suggested gestures are included
5. If two gestures are overall ranked equally, the one with the higher frequency among the older age group is selected

For each gesture selected for a particular task, a short video clip (2-5 sec) was recorded (at VGA resolution) that showed the execution of the gesture on top of the paper mockup for this task. Illumination, contrast and field of view were held constant across the video clips. The clips were edited to include a 1 second still at the end that visually retraced the gesture patterns by means of a superimposed red arrow, serving as memory aid after the clip had been played. The video clips always showed the mockup being held in the left hand while the right hand performed the gesture on top of a paper screen.

The 136 gestures selected for reevaluation using the above algorithm comprised manipulative gestures (53.7%), indexical gestures (21.3%), iconic gestures (18.4%) and symbolic gestures (alphanumeric and other symbols, 6.6%). None of the combined gestures made the selection, nor did any gesture that demands more than two fingers. The vast majority involved only one-finger (94.9%), and only 5.1% were two-finger gestures. The gestures consisted on average of 1.22 strokes ($SD=.56$, $Max=4$)

and 1.39 segments ($SD=.622$, $Max=4$). Complexity scores in this gesture set ranged from 1 to 6 ($M=1.73$, $SD=.94$, cf. Section 6.4.6).

6.8.2 Participants

The study was conducted with 42 participants, among them 22 younger (20-34 years, $M=26.1$ years, $SD=3.5$) and 20 older adults (62-74 years, $M=67.0$ years, $SD=3.9$). None of them had participated in study 2A before. The younger group consisted of 10 female and 12 male participants; the older group was gender balanced. Only one participant in the older age group indicated a slight physical restriction of hand or finger movement, and another one reported slight vision problems, which were crosschecked not to influence the outcome of this study. All other participants had normal or corrected-to-normal vision and no movement restrictions. All but one of the older participants and 86% of the younger participants were right handed. 85% of the older group were already retired, while 86% of the younger participants were students or employed. All participants had been using a mobile phone before, and the age groups did not differ in their years of mobile phone usage ($M_{young}=9.5$ years, $SD_{young}=2.24$ years; $M_{old}=7.76$ years, $SD_{old}=5.1$ years). Participants received a small financial reward for their participation.

6.8.3 Apparatus and Procedure

Participants were seated in front of two 17" TFT monitors (**Figure 6-8**). On the upper monitor ("task monitor"), the 34 tasks were presented sequentially, showing for each task a *before* screen (on the left), and an *after* screen (on the right), similar to the instructions given in study 2A. The goal for the participants was to find a suitable gesture, which transfers the current (*before*) state to the goal (*after*) state. To this end, on the lower touchscreen monitor ("gesture monitor"), which was mounted at an oblique angle (30°), four different gesture solutions for the task at hand were presented in the form of short video clips (see Section 6.8.1). Participants could playback the videos by tapping on them. Before activation, each video clip showed only the paper mockup with the task depiction in the *before* state. After playback, the arrow red arrow retracing the movement pattern remained visible.

Figure 6-8. Experimental setup: task descriptions are displayed on the upper monitor, gesture videos on the lower

At the beginning of each experimental session, participants were shown the basic interaction capabilities of an iPod Touch device, in order to demonstrate touch- and gesture interaction (without priming for any tasks that were to be tested). A training task (cf. Section 6.4.3) made them familiar with the paradigm, the touchscreen functionality and the solution space of possible finger gestures. The gestures shown in the training task included one manipulative, two symbolic (alphanumeric) gestures, and one iconic gesture; one of them was a two-finger gesture.

For each of the 34 tasks that followed, participants were instructed to carefully watch the four suggested gesture solutions on the gesture monitor and mark the perceived suitability for each of them on a 7-point Likert scale (1= not suitable at all, 7=extremely suitable). The procedure was self-paced and participants were allowed to watch videos several times, if desired. Additionally, they were handed a switched off iPod Touch that they could use to perform the gestures with their own fingers before they rated them.

The order of the tasks was varied in two versions, counterbalanced across participants, to minimize order effects. The position of a certain gesture solution on the gesture monitor was also counterbalanced across participants, so that each gesture had equal probabilities to appear in each quadrant of the screen. This was done to avoid any bias from the sequence the videos were watched.

After the rating procedure of the individual gestures, participants had to indicate which general attributes of gestures they generally perceive as important when judging the suitability of a gesture. The attributes were assessed with a short custom questionnaire with one item per attribute, and comprised *memorability*, *intuitiveness*,

complexity, efficiency, error tolerance, strain, uniqueness, number of fingers, posture, familiarity and curvature. Importance had to be stated on a 5-point Likert scale (1=not important, 5=very important). In addition, demographic data and prior experience with mobile phones, touchscreen systems and gesture control were assessed. The whole session lasted 1- 1.5 hours.

6.9 Results Study 2B

6.9.1 Gesture Evaluation across Age Groups

Older participants, across all gesture evaluations, rated gestures generally more suitable ($M=4.45$, $SD=.036$) than the younger users ($M=3.99$, $SD=0.5$; $t(40)=-3.307$, $p=.002$, $r=.46$). Since the verbal anchors of the suitability rating were chosen to be approximately equidistant, the rating scores will be treated at interval level (Rohrmann, 1978). Furthermore, for each participant, average rating scores were calculated from all gestures that shared a certain attribute, such as gesture *type*, *complexity score* or *number of fingers*. An analysis of variance with AGE GROUP (older, younger) as between factor and gesture TYPE as within factor (indexical, manipulative, iconic and symbolic gestures) revealed not only a main effect of age group $F(1,40)=16.090$, $p<.001$, $\eta^2=.287$), but also a main effect of gesture type $F(1,53,61.49)=43.031$, $p<.001$, $\eta^2=.518$), but no significant interaction. In general, indexical gestures (mostly simple taps) were found to be slightly more suitable ($M=4.66$, $SD=0.66$) than manipulative gestures ($M=4.5$, $SD=0.49$), and considerably less suitable were symbolic gestures ($M=3.33$, $SD=1.29$) and iconic gestures ($M=3.21$, $SD=0.92$).

Comparisons within each gesture type showed that the two age groups differed significantly in their suitability rating for indexical gestures ($t(40)=-2.16$, $p=.037$, $r=.32$), iconic gesture ($t(40)=-4.01$, $p<.001$, $r=.54$) and symbolic gestures ($t(40)=-2.49$, $p=.017$, $r=.37$), but not for manipulative gestures (**Figure 6-9, left**).

Figure 6-9. *Left*: Perceived suitability ratings (1=not suitable at all, 7=extremely suitable) according to gesture type. *Right*: Perceived suitability rating according to gesture complexity score. Error bars denote 95% confidence intervals.

Another ANOVA with AGE GROUP as between factor and NUMBER OF FINGERS (one-finger vs. 2-finger gestures) as within factor showed no general difference between one- or two-finger gestures, but revealed an interesting interaction with age group ($F(1,40)=10.50$, $p=.002$, $\eta^2=.208$): While younger users generally rated two-finger gestures more suitable ($M=4.23$, $SD=1.0$) than one-finger gestures ($M=3.98$, $SD=.51$), the older group rated vice versa. For them, one-finger gestures ($M=4.49$, $SD=.35$) were superior to two-finger gestures ($M=3.91$, $SD=0.96$).

The complexity score of each gesture was determined as described in Section 6.5.2. Because of the low incidence of gestures with a complexity score of 4 or more (6 was maximum observed) these were combined into a single category, and subjected to a 2 (AGE GROUP) $\times$ 4 (COMPLEXITY SCORE) analysis of variance. Besides the recurring main effect of age group, there was a main effect of complexity ($F(1.83, 73.09)=48.19$, $p<.001$, $\eta^2=.546$) and an interaction effect of complexity with age group ($F(1.83, 73.09)=48.19$, $p=.039$, $\eta^2=.081$). Both groups rated more complex gestures consequently as less suitable. However, the younger group punished increasing complexity more radically than the older group (**Figure 6-9, right**).

6.9.2 Agreement

For each task, gestures were ranked from 1 to 4 according to their suitability rating. Tied ranks were given for equal rating scores. The frequency of first ranks was counted for each gesture, and on this basis an agreement score was calculated for each age group, indicating the consensus of the age group that a best gesture match exists for a particular task. In **Table 6-4** and **Table 6-5** the best rated gestures are listed for each age group, together with the agreement score.

In general, and in accordance with study 2A, the younger group exhibited a significantly higher degree of consensus ($A=.376$) than the older group ($A=.320$), $t(33)=4.038$, $p<.001$, $r=.58$. **Figure 6-10** shows that agreement on the most suitable gesture is low (below .5) for virtually all tested tasks, and sometimes only slightly above 0.25, which represents a uniform distribution of “best gestures” over all four suggestions.

Table 6-4. Most suitable rated gestures for tasks in which older and younger share the favorite gesture. Sorted by suitability rating of older group; M = mean rating; SD = standard deviation; A =Agreement score

Task	Gesture	older			younger		
		M	SD	A	M	SD	A
Select multiple (list)	Swipe down across multiple objects	6.00	0.86	0.372	5.91	1.69	0.429
Select single	Tap on object	5.85	1.09	0.337	5.91	1.69	0.429
Close	Tap on object	5.55	1.36	0.311	5.85	1.04	0.413
Help	Tap on object	5.55	1.50	0.460	4.36	2.08	0.312
Play	Tap on object	5.55	1.19	0.319	5.73	1.12	0.441
Take call	Tap on object	5.55	1.19	0.295	5.68	1.58	0.327
Pause	Tap on object	5.40	1.23	0.308	5.73	1.12	0.441
Pan (scroll side)	Swipe left	5.35	1.49	0.388	6.45	0.80	0.506
End call	Tap on object	5.35	1.53	0.269	4.91	1.77	0.264
Select text	Swipe right across multiple objects	5.32	1.34	0.398	5.36	1.81	0.436
Rewind	Swipe left on object	5.30	1.42	0.342	6.09	0.81	0.440
Delete	Swipe right across object	5.25	1.55	0.308	4.59	1.76	0.272
Move object	Swipe up from object to another object	5.20	1.01	0.267	6.05	0.99	0.446
Select multiple (grid)	Swipe right, then down, then right across multiple objects	5.20	1.47	0.273	5.18	1.56	0.269
Fast forward	Swipe right on object	5.20	1.64	0.327	6.23	0.87	0.422
Stop	Tap on object	4.95	1.36	0.308	4.82	1.84	0.297
Write message	Swipe right across object	4.75	1.62	0.273	4.05	1.81	0.331
Open menu	Press & hold on object	4.55	1.70	0.251	5.50	1.65	0.421

Table 6-5. Most suitable rated gestures for tasks in which older and younger disagree on the best gesture. Sorted by suitability rating of older group; *M*= mean rating; *SD*= standard deviation; A=Agreement score

Task	older				younger			
	Gesture	<i>M</i>	<i>SD</i>	A	Gesture	<i>M</i>	<i>SD</i>	A
Make call	Tap on object	6.05	0.89	0.379	Double tap on object	5.50	1.23	0.491
Open	Tap on object	5.65	1.46	0.334	Double tap ob object	6.18	0.96	0.470
Previous page	Swipe left	5.55	1.47	0.380	Swipe left on object	5.23	1.85	0.333
Send message	Tap on object	5.47	1.39	0.284	Double tap on object	4.14	2.17	0.228
Scroll up	Swipe up across multiple objects	5.40	1.31	0.393	Swipe down across multiple objects	5.18	2.13	0.375
Next title	Swipe right on object	5.32	1.20	0.376	Number (title)	4.05	2.06	0.272
Confirm	Tap on object	5.30	1.49	0.281	Check mark on object	5.32	1.25	0.334
Volume up	Swipe up from object	5.25	1.33	0.288	Tap on object, then swipe up	5.50	1.23	0.357
Scroll down	Swipe down across multiple objects	5.20	1.67	0.314	Swipe up across multiple objects	5.14	2.21	0.430
Next page	Swipe right	5.20	1.39	0.346	Swipe right on object	5.27	1.80	0.328
Rotate	Circle clockwise	5.10	1.25	0.30	Press & hold while circle with 2 nd finger	5.14	2.08	0.359
Zoom in	Tap	5.00	1.65	0.288	Spread diagonal	5.67	1.77	0.435
Volume down	Swipe down to object	4.90	1.21	0.266	Tap on object, then swipe down	5.50	1.30	0.411
Zoom out	Tap	4.65	1.69	0.267	Pinch diagonal	4.77	2.58	0.444
Previous title	Arrow left	4.63	1.42	0.351	Number (title)	3.95	1.89	0.323
Cancel	Cross on object	3.85	1.31	0.269	Tap on neutral space	4.45	2.17	0.306

Figure 6-10. Agreement scores for different tasks and age groups

Only in 18 out of 34 (53%) tasks, the same gesture was rated as most suitable by the younger and the older group (**Table 6-4**). Not only does this result underline the difficulty of matching gestures to tasks which are intuitive to a vast majority of users in general, it also points to the differential judgment of suitability between the younger and older user groups.

Pairwise comparisons for each individual gesture (cf. **Appendix A.5.2**) reveal a significant difference in the rating scores between age groups in 49 out of 136 cases (36%). Breaking up the gestures that were rated significantly different by gesture type reveals that manipulative gestures were rated better similarly often by older users (13) than by younger users (10). However, indexical gestures (10 vs. 1), iconic gestures (12 vs. 0) and symbolic gestures (3 vs. 0) were clearly judged more suitable by the older group than by the younger group.

6.9.3 Persisting and Differing Effects across the Studies

While study 2A was of a generative nature, aiming at gesture synthesis, study 2B was of evaluative nature, aiming at gesture analysis. Despite the different approaches, some outcomes persist. Older users had suggested significantly more iconic and symbolic gestures, and these types of gestures were also rated significantly better by older users in study 2B. However, manipulative gestures, which had been suggested

more frequently by the younger group, were not found to differ along their perceived suitability. This type of gestures (e.g. swipes, circles, double taps etc.) did not spring to mind as often in the older group as in the younger group, but when presented, achieved similar preference levels.

Gesture complexity was also found to impact the different age groups differentially. In study 2A, older adults suggested gestures that showed a significantly higher level of complexity, and evaluated gestures of increased complexity less detrimental than the younger group. The same observation can be made for the number of fingers involved. Younger adults suggested significantly more multi-finger gestures in study 2A, and also rated them better than one-finger gestures in study 2B, whereas older adults initially suggested more single-finger gestures and rated those better than one-finger gestures in study 2B. Furthermore, in both studies the older group agreed less on the suitability of a particular gesture for a particular task.

This being said, a slightly different picture emerges if only the best-rated gestures are regarded. The total gesture set under evaluation consisted of 53.7% manipulative, 21.3% indexical, 18.4% iconic and 6.6% symbolic gestures. The older participants did not rate a single symbolic gesture as the most appropriate, and the fraction of iconic gestures was reduced to merely 5.8% (*arrow left* and *cross on object*). On the other hand, the amount of indexical gestures (almost exclusively simple tap gestures) was found to be disproportionately high (41.2%). The younger group was found to include at least two symbolic gestures among their favorites (*drawing the number of a title to switch back and forth between songs*), but from the 18% iconic gestures in the gesture set, only one was rated most useful (*check mark to confirm*). On the contrary, both indexical (26.5%) and manipulative gestures (64.7%) were disproportionately often selected as winners (**Figure 6-11**).

Figure 6-11. Fractions of different gesture types for the total set and for the subsets that were rated best in the younger and older group.

It is furthermore of interest to see how stable the age groups are in their preference ratings across the two studies. In other words, do the members of one age group rank those gestures best which were also ranked best in study 2A based on the frequency of their suggestions, or do they switch to gestures which were included in the gesture set based on the good ranks within the other age group? To this end, all gestures were ranked (1-4) according to their frequency of suggestion in study 2A and according to their suitability rating in study 2B in the respective age groups. The two ranks were correlated using Kendall's tau (because of the high number of tied ranks).

In the younger group, the gesture's rank in study 2B is highly correlated to its rank in study 2A ($\tau=.46, p<.001$), but not to the rank ordering of the older group in study 2A ($\tau=.10, p=.149$). However, the older group shows equal correlations to both older and younger gesture rankings from study 2A ($\tau=.28, p<.001$). This result can be interpreted such that the younger participants remained relatively stable (they rated those gestures best which they also suggested most frequently before), while the older group seemed to have changed their preferences slightly and now favor also some gestures initially proposed by the younger participants. Support for this interpretation comes from the fact that the total rank of a gesture in study 2B is higher correlated to the rank previously assigned within the younger group ($\tau=.41, p<.001$) than to the rank assigned within the older group ($\tau=.25, p<.001$).

6.9.4 General Criteria for Suitable Gestures

Participants were asked to state how important they regard certain criteria when deciding whether a gesture is suitable or not. The following general criteria were evaluated:

- The gesture is easily memorized (memorability)
- The gesture is understood intuitively (intuitiveness)
- The gesture consists of a minimum number of strokes (simplicity)
- The gesture can be executed particularly fast (efficiency)
- The gesture would work even in case of imprecise execution (error tolerance)
- The gesture can be drawn effortlessly (effortless)
- The gesture is distinctive (uniqueness)
- The gesture can be executed with only one finger (one finger)
- The gesture can be executed single-handed (single-handed)
- The gesture represents a familiar sign (familiarity)
- The gesture does not contain curves (linearity)

The single items were not systematically derived and not tested for their discriminatory power. Consequently, some criteria showed medium correlations, e.g. intuitiveness was negatively correlated to uniqueness ($r=-.36, p<.05$), posture correlated with

efficiency ($r=-.32, p <.05$), and linearity correlated with low simplicity ($r=.31, p <.05$) as well as error tolerance ($r=.49, p <.01$). Since no tradeoffs had to be made, most criteria were judged as important or very important. Good memorability and high intuitiveness were regarded as the most important features of good gestures, while familiarity of the pattern and a lack of curvature were seen least important (**Figure 6-12**). No differences in the importance of certain gesture aspects could be found between the age groups, except the criteria of familiarity, which was attributed a significantly higher importance by the older participants, $t(40)=-3.081, p=.004, r=.44$).

Figure 6-12. Mean importance ratings for several gesture criteria. (1= not important; 5=very important).

6.10 Discussion Study 2B

This present study was conducted with three major aims: First, to validate the results from study 2A, and find to out which of the previously found results are strongly related to the chosen paradigm, and which ones persist in a different paradigm. Second, to shed more light onto gesture attributes, gesture types, or even specific gesture patterns which show particular age-dependencies. Last but not least, based on the empirical data, to derive a gesture set which appears to be particularly suitable for older users, without compromising the suitability for younger users, and which can guide the implementation of a functional prototype.

6.10.1 Consistency with Study 2A

Regarding the consistency with study 2A, mixed results were obtained. In both studies, older users showed more affinity for iconic and symbolic gestures compared with younger users. Two related aspects might explain their stronger affinity for these types of gestures. The first reason is familiarity. Both iconic and symbolic gestures rely on the reuse of familiar signs. For example, drawing two parallel vertical lines resembles the *pause sign* (||) known from other music playback technology that participants are familiar with, or drawing the *check mark* gesture (✓) which can be found in many different contexts. Furthermore, in the questionnaire on general gesture criteria, familiarity was the only attribute that older adults marked significantly more important than younger users. This supports the hypothesis that familiar signs, such as found in iconic and symbolic gestures, are more appealing to older users compared with younger users. In fact, it is exactly this transfer of prior knowledge which can help to make a gesture intuitive to use (see Section 6.1).

The second reason can be found in the “mental model” underlying the gesture interaction. The effect of iconic and symbolic gestures is assumed only to trigger upon completion of the pattern. An automated gesture recognition system would have to wait, e.g. after the user has performed a downward swipe, whether another vertical swipe occurs (*pause sign*), another horizontal swipe occurs (*plus sign*), or the gesture is eventually continued with a horizontal swipe (e.g. *letter L*). In other words, gestures are understood as *commands*, which take effect only upon completion, similar to a button press, or to the selection of a menu entry. Manipulative gestures, in most cases, already take effect *while* the finger is moving. Lists are scrolled *while* the finger swipes down- or upwards, items are rotated *while* the finger performs a circular movement on the screen, etc. The older participants in our studies, even though they had been briefly familiarized with the direct interaction capabilities of a multi-touch device, might have stuck predominantly with the mental model of gestures as commands. The younger group, which generally had higher levels of exposure to multi-touch-enabled devices, was presumably more acquainted with the immediate nature of gesture interaction, and therefore rated gestures less suitable where they could not imagine a direct effect.

Also consistent with the study 2A is the divergent scrolling behavior in study 2B. The gesture rated best for both scrolling directions by the older participants was a swipe in the opposite direction compared with the gesture rated best in the younger group. As already discussed in Section 6.6.1, a different frame of reference, or “mental model” can account for these differences, which in turn can be traced back to different levels or prior experience with touch-enabled devices between the age-groups. Indeed, if correlations are calculated between prior gesture experience and the alternative gestures of swiping up or down in order to scroll down, or up, respectively,

we find positive correlations for a swiping movement against the intended scrolling direction ($r_{s_down}=.37$, $p<.05$, $r_{s_up}=.35$, $p<.05$) and negative correlations for swiping movements in the same direction as the scrolling ($r_{s_down}=-.35$, $p<.05$, $r_{s_up}=-.29$, $p=.059$). Partial correlation analysis further reveals that the significant correlations of scrolling direction with age vanish if gesture experience is accounted for. These results give further evidence that the preferred scrolling direction is habit-dependent and strongly depended on prior experience with similar or different technical systems, rather than a general preference of age group. Interestingly, this effect is only observed for the tasks of vertical scrolling. In the horizontal scrolling task (panning), both age groups agreed not only on the direction of movement, but also on the same gesture, since this task showed the highest agreement score of all. An explanation for this divergence can be found in the abundance of vertical scrollbars found in traditional window-based computing, which are believed to have influenced participants to apply the *window* metaphor rather than the *paper* metaphor, in contrast to the very rare occurrences of horizontal scrollbars.

The preference of one-finger over two-finger gestures among the older user group was substantiated in study 2B. The interaction effect between *number of fingers* and *age group* showed that younger and older adults perceive the usefulness of multi-finger gestures differently. However, the use of multi-finger gestures seems to depend strongly on prior experience with such gestures (Section 6.5.5) and should not be regarded as primarily age-related. The empirical data indicates therefore that multi-finger gestures are less suitable for novice users, while multi-finger patterns are more appropriate for more experienced users. However, since prior technology experience and expertise are usually negatively correlated with age, it is safe to say that for the majority of older users, simple one-finger gestures should be given preference by the system designer.

6.10.2 Divergence from Study 2A

Some results from study 2B differ from those obtained in study 2A. Most notably, the types of gestures that older users rated best in the multiple choice paradigm of study 2B differ from the ones they suggested most frequently in study 2A. For example, only in 20 out of 34 tasks (58.8%), the most frequently suggested gesture coincided with the gesture that was perceived as most suitable. Particularly, the iconic and symbolic gestures which had been suggested frequently by older users in study 2A (more than 1/3 were iconic and symbolic gestures; cf. **Figure 6-5**), no longer appeared in the list of most suitable gestures. Despite their generally higher affinity for these types of gestures, some of the initial proposals must be regarded as artifacts induced by the generative and ideative nature of the study. An example is the most frequently suggested icon, the *arrow* gesture. Its recurring appearance is probably

more an expression of the older participants' lack of understanding or imagination how direct manipulation can be effectuated on a multi-touch device, than an expression of preference for familiar and recognizable signs. When, in study 2B, a choice between an *arrow* gesture, and a much simpler *swipe* gesture was possible (effectively the arrow gesture without the arrowhead), the swipe gesture was rated more suitable (with one exception, the *arrow left* for previous title). Similar observations were made in most cases where alphanumeric symbols could be entered: while these gestures had been predominantly suggested by the older adults, none of them made it in their list of most suitable gestures in study 2B.

Agreement scores in study 2B were considerably higher ($A_{\text{young}}=.376$, $A_{\text{old}}=.320$) as compared to study 2A ($A_{\text{young}}=.247$, $A_{\text{old}}=.137$). This is, however, not surprising. In study 2A, agreement was calculated between gestures that were generated by individual participants, resulting in an average of 22.6 different alternative gestures for a task, many of which were not familiar to the other participants. In study 2B, on the other hand, agreement was calculated using only the four alternatives for a particular task known to all participants. The fact that even in the controlled multiple choice paradigm agreement scores were closer to a total disagreement ($A=0.25$) than to a total agreement ($A=1$) underlines the general difficulty of finding suitable gestures to a particular task. Even among the younger participants, who had all considerable computer, mobile phone- and touchscreen experience, and of whom all but two were familiar with at least the basic interaction concept of the iPhone, agreement never exceeded 0.506.

Two implications can be drawn from these observations. First, the gestures suggested by manufacturers of multi-touch devices for specific tasks need not be the only suitable and the only intuitive solutions; alternatives which could be equally suitable exist. Second, even for a particular task and a particular context, the idea of a single most appropriate gesture for all users might have to be given up. Different users with different levels of prior experience with varying technologies have different expectations, use different analogies, metaphors and frames of reference. System designers must carefully weigh the benefits of singling out one particular gesture for a certain task (e.g. in terms of learnability of the system, or overall consistency of system components) against a customizable gesture set, which can be adjusted to match the users' individual preferences.

6.10.3 Age Group Bias

The older participants rated the proposed gestures on average as more suitable than the younger participants. Since the gesture set was derived from the best proposals of both groups to equal amounts, it is unlikely that the gesture set in total fitted

better to the skills and knowledge of the older user group. Rather, older users tended to avoid expressing that a gesture was little or not suitable in contrast to the younger group, where a much wider spectrum of ratings was observed. Especially gestures that were rather complex received ratings between “not suitable” and “little suitable” from the young participants, while the older participants still perceived them as “moderately suitable”. One reading of these results is that older adults do not care as much about efficient interaction as younger users. Such a reading would be in accordance with their slower execution performance reported in chapter 5, and the findings by Sülzenbrück and colleagues (2010). However, if the list of best rated gestures in study 2B is regarded, the older group even showed an increase in the generally more efficient indexical and manipulative gestures compared with the younger group. A closer look at the best rated gestures of the older group reveals that 16 gestures involved swipes, 14 gestures involved a tap and only 4 involved another gestures. Since simple swipes and taps are among the most basic and efficient gesture patterns, another explanation becomes more likely. The generally less differentiated and more forgiving ratings of the older adults, in combination with their little differentiated selection of most suitable gestures, could be interpreted also in terms of a lack of imagination how a real gesture-controlled device would function, or as an expression of their heightened anxiety towards such technology and their resulting insecurity. Czaja and Sharit (1998) found that older people reported less comfort and less competence with computers, and also felt less control, and that these effects persisted even if prior computer experience was controlled for. Jay and Willis (1992), however, reported earlier that older adults’ attitudes towards technology can be altered, e.g. that direct experience with computers increased feelings of competence and comfort. It is thus highly probable that the perception of a gesture’s suitability will change once the older users gain some practical experience with a gesture-controlled device.

The list of most suitable gestures (**Tables 6-4** and **6-5**) should therefore be treated with care and not be implemented without further consideration. The present study cannot identify the criteria upon which suitability ratings were given. The high similarity between the gestures that were rated most suitable in study 2B (mostly swipes and taps) suggests that older users often opted for the most parsimonious solution, without considering a particular match of gesture and task. This interpretation is supported also by an observed lack of consistency and the detection of conflicts. A lack of consistency was observed in cases where antagonistic tasks (e.g. next title, previous title) show unrelated gestures (e.g. *swipe right* vs. *arrow left*). Conflicts appeared in those cases where the same gesture was considered best for different, sometimes antagonistic tasks on the same screen. For example, within the older group, *swipe down across multiple objects* was rated most appropriate for scrolling down a list, as well as for marking several items in a list; *tap on object* was chosen for

both open and close, and a simple *tap* was rated most suitable for zooming in as well as zooming out.

In conclusion, the results show that while there are considerable differences between age groups regarding the question which gestures are most suitable for a certain task, it still proves difficult to establish a gesture set which is particularly suited to older users based on mere empirical evidence from paper mockup studies. More insight can only be obtained from studies where gestures have actual effect. Thus, a functional prototype was devised and tested in study 3. The best-rated gesture set obtained in study 2B, however, must be revised to address consistency and conflicting gestures, before being implemented in the prototype.

6.10.4 A Revised Gesture Set

Even though participants had been instructed to consider the suitability of a gesture in context and to aim for consistency, the results show that this task was too difficult. Gestures were generally rated individually, and particularly so in the older group. The gesture set for older users was therefore revised with three major goals: First, conflicts needed to be resolved whenever the same gesture was suggested for different tasks on the same screen. Second, consistency needed to be ensured, such that the same (or highly similar) tasks on different screens are served with the same gesture, and that antagonistic tasks should be served, if possible, with antagonistic gestures. Third, while respecting goals one and two, the original preference derived from the user ratings should be maintained to the largest possible extent. If a gesture was identified to be conflicting or inconsistent, the second best ranked gesture was considered. The whole process was iterated until all conflicts and inconsistencies were resolved. The final revised gesture set is listed in **Table 6-6**.

Table 6-6. Revised gesture set for older users. Conflicts and inconsistencies resolved.

screen	task	gesture description
grid	select multiple grid	swipe right, then down, then left across multiple objects
list	select single	Tap on object
list	select multiple list	swipe diagonal down
list	move object	curve up from object to another object
list	delete	swipe right
list	scroll up	swipe up across multiple objects
list	scroll down	swipe down across multiple objects
list	open menu	Press and hold

list	play	Double tap
dialogue	confirm	Tap on object
dialogue	cancel	cross on object
closed folder	open	Tap on object
open folder	close	Tap on object
browser	next page	Swipe right
browser	previous page	Swipe left
map	zoom in	Tap
map	zoom out	Press and hold
map	pan	Swipe left (right, up, down)
map	rotate	circle clockwise
text	select text	Swipe (right) across multiple objects
homescreen	help	tap on object
music	fast forward	swipe right on object / tap on object
music	rewind	swipe left on object / tap on object
music	pause	Tap on object
music	stop	Press and hold
music	next title	Arrowhead right
music	previous title	Arrowhead left
music	volume up	Swipe up from object
music	volume down	Swipe down from object
incoming call	take call	Tap on object
calling	end call	Tap on object
contact list	make call	Double tap
contact list	write message	Letter M / Letter N
short message	send message	Double tap on object

6.11 Conclusion: Intuitive Gestures for Older Users

This second stage of the project was concerned primarily with the following questions: which gestures are particularly *intuitive* for older users, do these gestures differ from gestures for younger users, and, if so, on which dimensions? To this end, the notion of intuitive use was clarified to refer to *unconscious application of prior knowledge* which leads to *fast and effective* interaction, along with the definitions by

Blackler and colleagues (2002) and Mohs and colleagues (2006). Gestures as a means of interfacing technology were argued to be good candidates for facilitating intuitive use, since they can tap into deeper, more broadly shared knowledge (such as cultural or sensorimotor knowledge, Hurtienne & Blessing, 2007; Naumann, et al., 2007), and thus achieve intuitive interaction even for users that are less experienced with a certain product or technology domain.

A participatory design approach generated a wealth of gesture suggestions for typical mobile tasks. In order to classify the obtained gestures, a new gesture description language was devised with specification level necessary to describe gestures for recognition and synthesis, using a formalized grammar, but at the same time maintaining natural language to be understood across disciplines.

The results of studies 2A and 2B showed a strong preference for manipulative and indexical gestures across the age groups. Swipes and simple taps were the most frequently observed and also the best rated gestures. Nevertheless, studies 2A and 2B pointed at the general difficulty of finding gesture patterns which are perceived suitable by the majority of users, since individual preferences varied a lot and agreement was generally very low. Furthermore, the abundance of simple taps in the set of best-rated gestures, in combination with the low general agreement on a specific gesture for a task, suggests that not for all tasks a truly fitting and intuitive gesture exists. For some tasks, e.g. the ones on which little agreement existed, an interface solution relying on context menus and soft buttons is potentially more appropriate.

Concerning differences between younger and older user groups, studies 2A and 2B have shown that older users are more willing to trade efficiency for familiarity, and as such are more likely to accept symbolic and iconic gestures than the younger group. Furthermore, they are more tolerant towards more complex gesture patterns, even though less open for multi-finger gestures. If a system designer had to decide between a gesture pattern involving multiple taps or simultaneous finger movements, and a familiar though slightly more complex sign, the present results indicate that the latter would be more appropriate for an older user. Particular problems were observed for the older users to adopt the paper metaphor for scrolling lists. The persistency of this problem throughout the study can have been caused by the lack of feedback resulting from the static paper prototype, as discussed in Section 6.6.1. For these types of actions, specific instructions, training, visual aids or alternative scrolling mechanisms should be considered.

The studies have also highlighted the strong influence of prior experience. Strong correlations were observed between the participants' level of prior experience with mobile phones, touchscreens or gesture-based devices and the types of gestures they suggested or favored. Thus, the question whether a gesture can be used intuitively cannot really be answered universally, but depends to large extent on the prior

experience of a user with a specific type of product, with a specific interaction paradigm, or with technology in general. Results also show that sources of prior knowledge might interfere with each other, and that intuitiveness of use does not automatically increase if broader levels of the knowledge continuum are referenced (cf. **Figure 6-1**). For example, the *spread to zoom* gesture can be intuitive to a person with prior experience with an iPhone or similar product (product knowledge); a *tap to zoom* gesture can be intuitive to a person who is strongly familiar with zooming on a PC, analogous to mouse clicks (domain knowledge); *plus sign to zoom* can be intuitive also for people who have no experience with smartphones or computers, but rely on cultural conventions that plus is commonly associated with more, bigger, larger (cultural knowledge). The studies have demonstrated that for people who are not familiar with multi-touch enabled mobile devices, popular patterns such as *spread to zoom* are not intuitive and not rated most suitable. On the other hand, establishing a gesture set consisting only of widely familiar symbols and icons for older users has also been shown a misconception, since the parsimony of a gesture pattern and its efficient execution are equally important to older users.

A solution to this predicament could be, if technically feasible, to accommodate for different levels of prior knowledge and reference frames. Some very basic gesture patterns could be implemented as defaults, while shortcuts could be used for example through multi-touch patterns for experience users, and symbolic gestures offered as alternatives for users that prefer learnability and memorability over efficiency. In the example above, one could conceive to implement zooming at the same time as continuous zooming using a 2-finger *spread*, as discrete zooming using a simple *tap*, and simultaneously as a *plus sign* gesture for an alternative way of discrete zooming. If technically not feasible, multi-touch devices designed for older users should at least provide alternative gestures to such involving two or more fingers.

This chapter also pointed out the strengths and limitations of a participatory design approach. On the one hand, a richness of gesture suggestions was obtained, and sometimes seemingly well-established patterns were put to question. On the other hand, since the users – and particularly the older user group – lack the expertise and technical understanding, the results should not be regarded as design directives, but rather as a source of inspiration that can shift a designer's view from technological constraints towards the users' conceptual model. As Alan Cooper put it, "there is a large and growing body of evidence that users don't know what they want, don't know that the medium is capable of delivering, and are not quite incapable of imagining something new, useful, desirable, or innovative [...] But the results of those interviews is to the design solution as grapes are to wine: raw materials that must be transformed by expertise into a palatable product. " (Cooper, 2011, p. 1). Conse-

quently, the final “best-guess” gesture set presented in Table 6-5 is inspired by the older adults’ suggestions, but revised to be conflict free and consistent.

Lastly, the chosen approach has shown the limits of evaluating gestures in isolation. A particular gesture might be perfectly intuitive to use for a given task (and user), but still cause confusion in a larger context, if it contradicts or imitates other gestures. Other researchers have argued that a gesture set needs to be evaluated as a whole (Mauney & Le Hong, 2010), which is true, but could not be realized in this context. Not only the older user group, but also the younger group is overtaxed to evaluate 136 gestures and at the same time ensure consistency and avoid errors. A functional prototype (Chapter 7) was thus developed to test a specific implementation of the gesture set in context, and to reveal dependencies and interrelations between individual gestures.

7 Gesture Pragmatics: A Comparison of Gesture and Key Control

Stages one and two of the research project (Chapters 5 and 6) were primarily concerned with gesture syntax and semantics. Studies 1A and 1B could neither substantiate any specific problems of gesture-based interfaces for older users related to motor constraints, nor for small screens, difficult posture or complex patterns. Studies 2A and 2B showed that older users are relatively tolerant towards more complex patterns and, compared with younger users, have a stronger affinity for iconic and symbolic gestures, provided those are familiar. Thus, from a conceptual point of view, hardly any constraints need to be assumed for a touch and gesture-based interface for older adults. On the contrary, between 70% (study 1B) and 90% (study 2B) of the older participants stated that they would like to use similar gesture-controlled systems in the future. On a cautionary note, however, in all four studies gestures had not been investigated on a functional device and in context.

Based on the encouraging results from studies 1A-2B and the list of potentially suitable gestures derived in Section 6.10.4, a functional prototype was developed and tested in study 3 to address pragmatics of gesture interaction: how well does gesture control work for older adults in a real interaction context? This chapter comprises the implementation stage (Section 7.1) and evaluation stage (Section 7.2 -7.4) of the project as described in Section 4.3.2.

The basic research question behind study 3 was: Does the touch-and gesture interaction provide any added value to older adults and, thus, facilitate technology interaction, or might it – on the contrary – constitute a new obstacle for technology adoption? In order to answer this question, a benefit or added value compared with existing interface paradigms needs to be quantified. To this end, the gesture interaction mode was compared with the traditional key control of a mobile device. Subjective and objective measures of usability were used to quantify advantages and disadvantages of gesture input for older users. The basic research question can be broken down into more specific questions:

1. Is gesture input more *effective* than conventional key input for older users?
2. Is gesture input more *efficient* than conventional key input for older users?
3. Can a gesture-controlled device be used more *intuitively* than a key-controlled device?

4. Can gesture commands be *memorized* better than key commands and related menu entries?
5. Is gesture input equally *accepted* by older adults as compared to conventional key input?

Success rate was not an option to quantify effectiveness, because all participants were urged to complete all tasks in order to have sufficient data points to quantify other interaction parameters. Instead, the *perceived achievement of goals* subscale of the QUESI questionnaire (Questionnaire for Intuitive Use, Hurtienne & Naumann, 2010) was used to provide an estimate. Efficiency was quantified through *time-on-task* and *interaction steps* until goal achievement, as well as the subjective *efficiency rating* from the SUMI questionnaire (Software Usability Measurement Inventory, Kirakowski, 1996). Since interaction errors were of very different nature in both experimental conditions (gesture vs. key control), the measure that lends itself best for comparison was to count *redundant screen changes*, and to rely on the subjectively *perceived error rate* (QUESI subscale). Intuitiveness was quantified through the global *intuitive use* score from the QUESI questionnaire. No measure was found for a fair comparison of memorability between the two input modes. Instead, the *recall rate* of gesture functionality was assessed. Finally, acceptance of a system is not only hard to define, but also hard to assess after a one-time interaction session. Measures of *affect* (SUMI subscale), *preference* and *intention to use* can, nevertheless, give a first indication how well the input paradigm is received among this age group beyond pure performance measures. The dependent variables are described in greater detail in Section 7.2.3.

Terminology: In the remainder of this chapter, the terms *gesture input*, *gesture control* and *gesture interaction* are used in a synonymous manner, short for gesture-based input/control/interaction. Similarly, the terms *key input*, *key control* and *key interaction* are used. Whenever the two experimental conditions are referenced, we speak of *gesture* and *key mode*. Notably, in contrast to other literature (e.g. M. Karam, Lee, Rose, Quek, & McCrickard, 2009; Naumann, Wechsung, & Hurtienne, 2010), we do not differentiate between touch and gesture control. Instead, in line with the definition of gesture (cf. Section 3.1.3), touches (i.e. simple taps) are treated as the most basic way of gesturing and are thus subsumed under *gesture control*. However, since the operating system of the prototype handles taps fundamentally differently from user-defined gesture patterns, they are sometimes described separately or not featured in certain analyses.

7.1 Development and Description of the Functional Prototype

The main goal during the development of the functional prototype was to create a system that

- allowed for both gesture and key control
- provided a realistic touch and feel
- was generic enough to allow transfer to similar devices
- was able to reliably detect custom gesture input
- incorporated a logging tool to record all user input

The testing device should ideally include a multi-touch enabled touchscreen, a 5-way directional keypad and an Android platform. The Android operating system was regarded as most suitable for custom software development.

7.1.1 System

Of fourteen current smartphones with keypads that we compared, none fulfilled all requirements. The Samsung Galaxy Spica i5700 was chosen as the best option since it provided the Android 2.1 platform, allowing to use the Android Gesture libraries, a keypad with relatively large keys, and a touchscreen, which, unfortunately, was restricted to single-touch only (**Figure 7-1**). However, this drawback seemed acceptable, since previous studies had demonstrated a preference for one-finger gestures among the older users (see Sections 6.5.2 and 6.9.1).

Figure 7-1. The functional prototype with a trace of a gesture

The functional prototype was developed from scratch during a period of six months based on Java™ and the Android 2.1 API. We employed the Eclipse IDE for Java Developers and the Android SDK as development tools.

The final experimental software consisted of 30 java classes which provided the interactive features of the UI, logging functionality, and the gesture detection functionalities. The layout of the different screens was determined through 39 XML files, one for each screen, tab or dialog. Almost two hundred images were used depicting icons, gesture patterns, photos, or other interface elements. The icons and pictures used in the test software were either custom designed or downloaded from the web, provided they were free to use under the *Creative Commons* or similar licenses.

7.1.2 Screen Design and Functionalities

The homescreen of the device was designed along three goals:

1. Giving easy access to functionalities that have been identified to be frequently used by elderly users.
2. Providing a high degree of realism by taking reference from existing home screen concepts.
3. Keeping consistency with the with previous studies by incorporating the investigated functionalities .

Previous research has pointed out that elderly users predominantly make use of the communicative features of their mobile phones, rarely going beyond making and receiving voice calls (Kurniawan, 2008). Renault and Bijon (2008) list phone book, SMS, dialing a number, alarm and camera among the most frequently used features of older adults. Similarly, a study by Kurniawan and colleagues (2006) revealed address book, SMS and alarm clock as the top three desired functions besides calling. A large scale survey conducted by the PEW Research Center (N=605 users aged 60 and above) showed that only the camera (43%) and text message (26%) functionalities have been used by a larger group of elderly users, while other functionalities including email (12.7%), internet (10.1%), instant message (10.1%), games (7.9%), music (7.1%) and video (6.3%) have been used only by a minor fraction (Pew Internet & American Life Project, 2010). Accordingly, the homescreen should contain only a limited number of items, providing easy access to dialing, address book, text messaging, and potentially also include camera and alarm clock icons.

An analysis of 20 different smartphones from different manufacturers and varying operation systems was conducted to reveal the most prominent functions and their most typical position. Based on frequency of occurrence, a typical smartphone homescreen features icons for address book (90%), internet (85%), calendar (80%), messages (80%), music/player (65%), media/pictures/gallery (55%), phone (50%), preferences (50%), call log (50%), camera (45%), alarm (40%), maps (35%), chat (35%) and email (30%). The surprisingly low prevalence of phone icons can be explained by

the fact that many devices still feature specialized hard keys to access the basic phone functions.

In order to incorporate most gesture patterns for tasks tested in study 2B (see **Table 6-6**), the prototype should include generic user interface (UI) elements like grids, lists, dialogs or textfields, and the specific functions map, browser, music, calling, contact list and text message.

Based on the findings from gerontechnological research, consistency with existing devices and constraints from the previous studies, the final homescreen design included icons for *contacts*, *message*, *calendar*, *phone*, *internet*, *music*, *gallery*, *clock* and *maps* (**Figure 7-2**). Even though the camera functionality had been found to be a commonly used feature also among older users, it was left aside because of its strong linkage to the hardware, and the potential difficulties to design a realistic mock-up. Since maps and music were not incorporated in the final set of tasks (Section 7.2.4), a mockup of these functionalities was not included in the prototype. The homescreen was devised as a square grid of nine icons, 72x72 Pixel each, following the recommendations on icon size, spacing and concreteness (Jin, et al., 2007; LaVictoire & Everhart, 2009; Schröder & Ziefle, 2006, 2008).

Figure 7-2. Homescreen of the test application. The incorporated functions are (from top left to bottom right): contacts, message, calendar, phone, internet, music, gallery, clock, and maps

The position of common homescreen elements varies a lot across devices. Different principles seem to apply: more important functions are placed for easier retrieval in the upper left corner while less important ones towards the lower right (e.g. Neisser, 1963; Somberg, 1986), but one can also find important items placed in the lower row to allow for easy thumb activation, or in the center to shorten key navigation.

7.2 Methods Study 3

7.2.1 The Gesture Set

Fourteen different gestures (besides taps) were implemented to control the device. They included six direct manipulation gestures and eight symbolic or iconic gestures. Only one gesture (*cross*) consisted of two segments (one stroke each), the rest were single-segment gestures. All gestures could be drawn using only one finger. **Table 7-1** shows the set of implemented gestures and their associated functions.

Table 7-1. Gesture set with associated functions. Arrows indicate the direction of finger movement

direct manipulation gestures			symbolic gestures
swipe up	<i>scroll down in lists</i>		arrowhead up	<i>return to homescreen</i>
swipe down	<i>scroll up in lists</i>		arrowhead left	<i>return to last screen</i>
swipe right	<i>scroll left</i> <i>highlight underlying text</i>		letter S	<i>Send item</i>
swipe left	<i>scroll right</i> <i>highlight underlying text</i>		letter L	<i>open contacts List</i>
curve up	<i>move item up</i>		letter N	<i>compose message</i> <i>(Nachricht)</i>
curve down	<i>move item down</i>		question mark	<i>open help screen</i>
			cross	<i>delete item</i>
			heart	<i>open favorites list</i>

The gesture set presented here differs substantially from the set described in Section 6.10.4. Reasons for this divergence are manifold:

- Technical constraints and limitations of the gesture recognition software in the device prohibited the use of certain gestures.

- Dialogue boxes with virtual buttons (e.g. to confirm or cancel) were giving preference over gestures (e.g. *check mark*) in certain situations after pretesting
- Additional navigation gestures needed to be provided, which were not part of the tested gesture inventory in study 2B (e.g. *help* and *return to home* or *previous screen*).
- To ensure that a gesture which has to be available on every screen (e.g. *help* or *return to home screen*) does not interfere with gestures for screen-specific functions
- The internet navigation task (simplified to work without a virtual keyboard) necessitated to introduce a gesture to call the list of favorites (*heart* gesture), which had not been evaluated in study 2B.
- The gesture recognition methods in the Android libraries used algorithms which favored symbolic gestures over continuous, direct manipulation gestures, resulting in a relatively high proportion of symbolic and iconic gestures in this gesture set, in contrast to the original list (Section 6.10.4).

Despite several changes, care was taken that, where possible, gestures still corresponded to the ones that were positively rated by the older users in study 2B, and that the gesture set itself was consistent and worked throughout the whole application. Due to the rare occurrence of iconic, alphanumeric and other symbolic sign gestures, these were summarized under the term *symbolic gestures* within the scope of this study in order to obtain meaningful quantifications.

7.2.2 Apparatus

The hardware on which the interaction tasks had to be carried out was the Samsung i5700 (or Samsung Galaxy Spica) smartphone, as described in Section 7.1.1. It featured a 3.2in (81mm) capacitive HVGA touchscreen with 320 x 480 pixels at 180 pixels per inch (ppi). The whole device measured 115mm x 57mm x 13.2mm (HxBxD) and weighed 124g. The operating system was Android 2.1. The device had a keypad below the touchscreen, consisting of 11 hardware keys. Two larger plus one smaller function keys were located at the sides, and in between a diamond of 4 navigation keys with an OK key in the center (**Figure 7-1**). Besides the navigation keys and the central confirmation key, only three of the large function keys were used in the study. The other keys were – if possible – overridden in the software to have no effect, or otherwise participants were instructed not to press them.

7.2.3 Design

The main experimental manipulation was the INPUT MODE: gesture vs. key interaction. Each participant first had to operate the smart phone for a series of tasks in one of the modes, and subsequently in the other mode. The ORDER of the two conditions was counterbalanced across participants and serves as a between factor in the analysis to check for any sequence effects. The independent variable TASK varied as within-factor on five levels. The five tasks (see Section 7.2.4) were repeated in each input mode condition in the same order, but varied slightly in two highly similar scenarios (A and B) to avoid repetition. The two scenarios were counterbalanced across input mode and order.

The efficiency and effectiveness of interaction was quantified with the dependent variables *time on task*, *redundant steps*, *steps percentage*, and *redundant screen changes*. *Time on task* is measured as the total time passing between the first and the last logged interaction step within a task.

As *interaction step* count all those events from the behavioral data in the logfile that clearly can be associated with user input. This could either be a tap or long-press on the touchscreen, a direct manipulation or symbolic gesture, or a key press. These also include interactions which had no direct effect, such as tapping the screen at non-functional areas of the GUI, or pressing a key with no associated function. For each task in both scenarios and input modes an optimal number of interaction steps exists, which is the minimal number of taps, gestures or key presses leading to task completion. These minimally required steps differ slightly between scenarios, but differ strongly between input modes (cf. **Appendix A.6.1**). In all cases, the key condition demanded for more interaction steps than the gesture condition. We measured the total number of interaction steps the participant needed to complete the task, and subtracted the optimal number, leaving only the excess or *redundant steps* as a measure of how efficiently the task was accomplished.

Besides the absolute number of interaction steps, the *steps percentage* measure puts the effectively needed interaction steps in relation to the optimal path which was set to 100%. A 200% *steps percentage* indicates that the number of interaction steps was twice as much as the optimal number in a specific input mode.

As another measure of task efficiency, the number of *redundant screen changes* was counted. A screen change occurred whenever the system had to load a new graphical layout. For example, sending a picture included four screens: homescreen, gallery grid, gallery and the picture-sending dialog, and thus three screen changes. As the optimal number of screen changes was the same in the key and gesture conditions, the number of unnecessary screen changes is an estimate of navigation efficiency.

The behavioral data was complemented by the questionnaire data from the SUMI (Kirakowski, 1996) and QUESI questionnaire (Hurtienne & Naumann, 2010). The SUMI is a psychometric tool to assess the usability of a system, consisting of 50 items, 10 for each of the subscales *efficiency*, *affect*, *helpfulness*, *control*, and *learnability*. The QUESI measures the subjective consequences of intuitive use with 14 items, and consists of the subscales *subjective mental workload*, *perceived achievement of goals*, *perceived effort of learning*, *familiarity* and *perceived error rate*.

7.2.4 Procedure

The experimental session began with a short demographic interview and introduction to the general background and purpose of this study. The participants then received written instructions related to the specific input mode they were going to use, introducing the basic functionalities and particularities of this mode (see **Appendix A.8**).

The instruction for the gesture condition comprised:

- description of the homescreen
- description of the interaction through finger press and finger movement
- example of the question mark gesture and its function to open the help screen
- hints where gestures work and how to best hit a virtual button with the fingertip

The instructions for the key condition comprised:

- description of the homescreen
- visual depiction of key functionalities (as depicted in **Figure 7-4**)
- tip to use a long-press on the OK key in certain situations

Figure 7-4. Operation instructions for the key mode. Keys needed for operation included *menu*, *OK*, *back*, *home* and *navigation keys*

In case of the gesture condition, participants received a brief training before they were given the actual tasks. During the training, all fourteen possible gesture patterns were depicted sequentially in the upper right corner of the screen, and the participants had to redraw the depicted pattern on an empty screen. A yellow trace visualized the performed movement. The application only advanced to the next pattern once the gesture was detected correctly by the software, requiring participants to redraw the pattern until they drew it correctly. This training familiarized the participants with the touch and feel of the device and the way it reacts to the finger movement input and gave them an overview of the gestures available for interaction.

In both conditions the participants were given one of two very similar scenarios (see **Appendix A.8**), each consisting of the following tasks:

1. Calling a contact from the contacts list
2. Editing and sending a message to a specified contact
3. Retrieving an entry in the calendar and moving it to a different time slot
4. Retrieving a specified picture from the image gallery and sending the picture to a specified contact
5. Retrieving and opening two different websites from the bookmarked favorites

Participants were asked to interact with the device in each condition until all tasks were accomplished. Help was only provided by the experimenter in clear situations of helplessness and in a strictly structured manner according to previously established helping guidelines.

After completing the tasks, participants were given two questionnaires (SUMI and QUESI, see Section 7.2.3) in order to rate the experienced interaction in terms of usability and intuitiveness. After the gesture-control condition, participants had to fill out an additional questionnaire which depicted all the possible gesture patterns. For each pattern they had to indicate whether they used it, remembered the associated functionality, and how they rated the aptness of the pattern. After completing both interaction conditions, a short interview probed for their preference regarding the interaction mode, the reasons for this preference, and improvement suggestions. At the end of the experimental session, participants had to fill out a custom-designed questionnaire related to their previous experience with mobile phones, touchscreens, gesture-control and iPhone-usage, and to complete the computer literacy scale (CLS, Sengpiel & Dittberner, 2008), a questionnaire assessing their computer proficiency.

7.2.5 Participants

The prerequisites for participation in this study were an age between 60 and 75, no major pathological restrictions in manual dexterity, no major impairments to vision, possession of a mobile phone and a minimum degree of expertise with the mobile

phone. This comprised the ability to dial a number, call a contact from the address book and send a short message (SMS). Prerequisites were checked through self-report during telephone interviews. In total, 35 (16 male, 19 female) German native speakers were invited to take part in this study.

Three participants were not able to finish the study within the given experimental conditions. They did not complete all tasks in both experimental conditions, because they had problems with key control. One participant was not able to grasp the concept of moving focus from item to item with the help of the navigation keys. The experimenter aborted the session when it became obvious that despite generous hints the tasks could not be achieved. Another participant did not understand that the navigation keys – although visually connected – were in fact four separate direction keys. The experiment had to be aborted after extensive help and the failure to reach beyond task two. The last participant to be excluded managed to complete all tasks in key mode, but it took her much longer (39 min) than the rest of the participants (10 – 30 min), rendering a completion of the whole experiment in time impossible. Since only incomplete data was available for these three participants they were excluded completely from all further descriptive and inferential statistics. Surprisingly, all dropouts are related to difficulties in the key mode, and all occurred when key control was the first condition.

The sample which underwent statistical analysis, therefore, consisted of 32 participants (gender-balanced), between 61 and 75 years old ($M=68.5$, $SD = 4.4$). Female participants ($M=66.9$, $SD=4.5$) showed an age distribution similar to male participants ($M=69.6$, $SD=3.9$). Two participants were left-handed, the rest was right-handed. All had normal or corrected-to-normal vision (on self-report), and only three reported minor handicaps of their manual dexterity, such as a reduced thumb flexion, reduced flexibility in the ring finger, or slight numbness in the fingers after extensive movements. The educational background of the participants was mixed, with 28% higher education, 12.5% high school and 44% secondary school completion. The vast majority of the participants were retired (87.5%), only 12.5% percent reported to work part-time or full-time.

7.2.6 Previous Experience

All participants were required to have at least a little previous experience with mobile phones. They usually used only a very narrow range of functions, most notably the basic phone functionality, followed by messaging and clock (see **Figure 7-5**). These patterns of use correspond to previous studies and surveys on older people's mobile phone usage habits (e.g. Kurniawan, 2008; Kurniawan, et al., 2006; Pew Internet & American Life Project, 2010; Renaud & van Biljon, 2008).

Figure 7-5. Frequency of use for common mobile phone functions

On average, participants had previous experiences with three different kinds of touchscreen devices or systems. Most of them were familiar with touch-operated ticket machines (85.7%), cash machines (82.9%) and info terminals (51.4%). Touchscreen experience of the older users came mostly from public stationary terminals with large screens, while personal mobile devices with smaller touch screens were far less familiar, for example navigation systems (31.4%), digital cameras (20%), mobile phones (17.1%) or mp3 players (2.9%). Only two participants reported never to have operated a device via touchscreen before.

Even though roughly one in six participants indicated to have used a touchscreen-operated mobile phone before, this might involve brief exploration episodes only, since 94% stated that their own current mobile phone was key-controlled. The exceptions were a 75-year-old lady who used an Apple iPhone, and a 66-year-old gentleman using a LG smartphone.

Assuming the Apple iPhone, was the most popular and widely known example of a device that allowed for finger gesture control, we explicitly asked the participants to which degree they were familiar with this product and its usage pattern: the majority of our sample either did not know this device at all (25%) or heard of it, but knew nothing about its way of using (62.5%). Only five participants knew about the basic usage patterns through observation or media, while - except for the one participant who owned one - only one other participant reported to have used this device himself at least once.

7.3 Results Study 3

The results were analyzed using repeated measures analyses of variance (ANOVA) assessing effects of INPUT MODE as within factor and ORDER as between factor on the measures of efficiency and the perceived usability (SUMI) and intuitiveness (QUESI) scales. A subsequent repeated measures ANOVA with TASK and INPUT MODE as within factors revealed the relative impact of the different tasks on the overall performance. In the repeated measurements analyses where the assumption of sphericity was violated, the Greenhouse-Geisser corrected values are reported. Effect sizes were calculated as partial eta square (η^2). An alpha level of .05 was used for all statistical tests.

7.3.1 Comparison of Input Mode

One major question driving this study was to investigate how well participants performed in the two interaction conditions, and in particular so, whether the presumably less familiar gesture interaction mode matches the presumably more familiar input mode with navigation and menu keys. Since there was no time pressure and all participants were brought to complete all tasks, measures of efficiency were of prime interest. The first thing to notice is that participants were on average 19% faster ($M=15$ min 51 sec; $SD=5$ min 55 sec) in the gesture mode as compared to the key mode ($M=19$ min 26sec; $SD=5$ min 50sec) ($F(1,30)=15.1$, $p=.001$, $\eta^2=.334$). There was no effect of the order in which participants had to operate the device, nor an interaction between input mode and order. The fact that the gesture condition was performed faster is in line with the significantly reduced number of redundant interaction steps that participants exhibited in this mode, $F(1,30)=55.7$, $p<.001$, $\eta^2=.65$). While in gesture mode the older users tapped or gestured without need on average 92.2 times ($SD=48.0$), they produced more than double the amount of redundant interaction steps in the key condition ($M=257.2$, $SD=134.0$). Whether they had already experienced the device in another interaction mode, or not, made no difference.

While this increase of redundant steps by 179% does not speak in favor of the key control in terms of efficiency, a more telling measure is the relative percentage of redundant steps, since it takes the different base rates of interaction steps for the two conditions into account. Even though participants were still more efficient in the gesture condition - they needed on average 283% of the optimal interaction steps ($SD=95.9\%$), compared with 293% ($SD=101.7\%$) using the keys - this difference was no longer significant. Again, there was no interaction with order.

Consistent with the measures of time and redundant interaction steps was the significant difference in redundant screen changes ($F(1,30)=55.513$, $p<.001$, $\eta^2=.65$). In the gesture condition participants deviated from the optimal path of screen successions on average 12.7 times ($SD=8.7$). When using the keys they were swapping the screens much more often ($M=30.57$, $SD=15.7$). This is the only measure of efficiency where we observed a significant main effect of the order of conditions ($F(1,30)=5.91$, $p=.021$, $\eta^2=.17$). Participants were generally less efficient ($M=25.9$ redundant changes) when they had to interact first in the gesture condition as opposed to performing first in the key condition ($M=17.3$ redundant screen changes). There is no general improvement in efficiency from the first to the second condition. Rather, it seems to be the case that first using the device with the keys forces them to explore the device extensively, including opening and re-opening redundant screens, resulting in a more straightforward accomplishment of tasks with less redundant screens in the second gesture interaction block. On the contrary, experiencing at first a smoother transition of tasks using gestures seems detrimental to a subsequent interaction with keys, where navigation through screens is more cumbersome (Figure 7-6).

Figure 7-6. Mean number of redundant screen changes for both input modes according to the order of interaction. Error bars denote a 95% confidence interval

Taken together, across all different tasks, participants were more efficient using touch and gesture than using the keypad, even though they had little to no previous experience with this interaction style on handheld devices.

The findings are in line with the data from the SUMI and QUESI questionnaires. For each of the subscales of the SUMI questionnaire, as well as the compound global usability score, a Wilcoxon signed-rank test showed a significant difference between the gesture and the key mode. Participants rated the *efficiency* of gesture control higher ($M=48.8$, $SD=13.6$) than key control ($M=42.6$, $SD=13.8$), $z=-2.142$, $p=.031$, $r=-.20$. Gesture control also achieved significantly better ratings on the subscales of *affect*

($z=-3.354$, $p<.001$, $r=-.42$), *helpfulness* ($z=-4.408$, $p<.001$, $r=-.55$), *control* ($z=-2.470$, $p=.012$, $r=-.31$), and *learnability* ($z=-2.920$, $p=.003$, $r=-.37$, **Figure 7-7**). Moreover, the global measure of usability calculated from the SUMI item scores shows the perceived advantage of the gesture interaction ($z=-3.451$, $p<.001$, $r=-.43$). Looking at the effect sizes, especially the better helpfulness and the users' better emotional reaction towards the system, or "likeability", contributed most to the overall better rating.

Figure 7-7. Mean scores for the SUMI questionnaire according to input mode. Error bars denote a 95% confidence interval

Gesture control was also rated more *intuitive* to use, as confirmed by the total score of the QUESI questionnaire ($z=-2.45$, $p=.013$, $r=-.31$). This QUESI data further tells us that participants considered the gesture version to allow for a higher degree of goal achievement ($z=-2.513$, $p=.011$, $r=-.31$), and a reduced learning effort ($z=-2.388$, $p=.015$, $r=-.30$) than the same system operated by keys. This is in agreement with the better learnability result from the SUMI questionnaire. The Wilcoxon signed-rank comparison for perceived error rate yields a marginally significant advantage for the gesture control ($z=-1.917$, $p=.055$, $r=-.24$). However, since differences between groups in this case can be assumed to be normally distributed (K-S test: $D(32)=0.14$, $p=0.98$), a dependent t-test gives us a stronger indication ($t(31)=-2.183$, $p=.037$, $r=.37$) that participants perceived their error rate in the keys condition as being higher ($M=2.55$, $SD=0.96$) than the gesture condition ($M=2.13$, $SD=0.97$).

While the original QUESI items were coded in a way that a higher score always relates to a higher degree of intuitive use (e.g. a higher average value in "perceived error rate" actually meant a higher degree of intuitive use on that subscale and thus a lower perceived error rate), we recoded the subscales for *perceived error rate*, *perceived effort of learning* and *subjective mental workload* for better readability. As depicted in

Figure 7-8, a higher average item score now corresponds to a higher error rate, higher learning effort and greater mental workload respectively. There were no significant differences in the subscales of subjective mental workload ($p=.21$) and familiarity ($p=.077$).

Figure 7-8. Mean scores for the QUESI questionnaire according to input mode. Error bars denote a 95% confidence interval

Taken together, the older users found the gesture control of the device in general more intuitive to use, because they thought they could achieve their goals better, they had the feeling to learn this system faster and perceived themselves to make fewer errors compared with the keys condition.

The analyses of efficiency from the logfile data as well as the user ratings from the SUMI and QUESI questionnaires show a general benefit of the gesture input as compared to the key input. This pattern is further substantiated by the older users' preference statement: when asked which of the two interaction modes they preferred, 72% stated they preferred the gesture condition, while only 22% uttered a preference for the keys condition and 6% were indifferent.

7.3.2 Task Effects

Across all tasks, the gesture condition showed a general advantage on several objective and subjective measures. However, a benefit might not occur for all tasks in equal measure, or the overall advantage might be traced back only to a very specific interaction situation where the keys were in comparison most cumbersome to use. We therefore conducted further repeated measures analyses of variance with TASK as 5-level within factor, besides the 2-level within factor INPUT MODE. The tasks

will be referred to as *calling*, *message*, *calendar*, *gallery* and *website* in the following (see Section 7.2.4 for details). Significant main effects of input mode that have already been reported in the previous section are omitted here; instead we focus on task effects and, most interestingly, the interaction of task and input type.

Time on task. The average time needed to accomplish a task did differ significantly from each other ($F(4, 124)=19.86, p<.001, \eta^2=.39$). The average time needed to successfully complete the first task – calling a person from the contact list – clearly was the shortest ($M=2:33$ min, $SD=1:12$ min), and the second task – editing and sending a text message – the longest ($M=4:51$ min, $SD=2:41$ min). The remaining tasks involving the calendar ($M=3:27$ min, $SD=1:52$ min), the gallery ($M=3:23$ min, $SD=1:51$ min) and opening two websites ($M=3:29$ min, $SD=1:47$ min), were comparable and needed a medium amount of time. Pairwise comparisons (Bonferroni corrected) show that the *calling*, and *message* tasks differ significantly from the other tasks, while there is no time difference between tasks 3, 4 and 5. As the tasks naturally varied in complexity, and required different numbers of interaction steps, these differences are not surprising. Interesting is the significant interaction between task and input mode ($F(2.9, 91.1)=4.031, p=.01, \eta^2=.115$), meaning that tasks were handled with diverging levels of efficiency between the gesture and the key control conditions. Comparing the mean times for each task separately with dependent t-tests revealed a significant difference in task times only for the *message* ($t(31)=-2,596, p=.014, r=.42$) and *calendar* task ($t(31)=-4.035, p<.001, r=.57$), while on the other tasks there was no statistical difference between key and gesture input. The message task was performed on average faster by 1:39 minutes, the calendar task by 1:45 minutes when using gesture input (Figure 7-9).

Figure 7-9. Completion times for different tasks.
Error bars denote a 95% confidence interval.

Redundant interaction steps. For the redundant interaction steps we found a significant main effect of task ($F(1.8, 55.5)=41.844, p<.001, \eta^2=.57$) as well as a significant interaction of task and input mode ($F(1.7, 52.4)=30.73, p<.001, \eta^2=.498$). In all tasks, the gesture condition produced a significantly lower amount of unnecessary interaction steps than the key condition (**Figure 7-10**). While for the tasks of *calling*, *gallery* and *website* the difference in redundant interaction steps between the operating conditions remains comparable (13.4, 19.4 and 15.6 steps respectively), it is again the *calendar* task (66.3 steps difference) and in particular the *message editing* task (133.3 steps difference) that account for a disproportionate increase in redundant interaction steps.

Figure 7-10. Redundant interaction steps.
Error bars denote a 95% confidence interval

Taking into account the different number of minimally needed interaction steps for these tasks (see **Table A.6-1**), a similar picture emerges. Again, we find an influence of task on the *percentage interaction steps* ($F(4, 124)=11.735, p<.001, \eta^2=.275$) and an interaction with the input mode, $F(3, 94.1)=3.774, p=.013, \eta^2=.109$). This measure confirms that tasks were performed with varying efficiency, with *message* ($M=211.5\%$, $SD=134.4\%$) and *calendar* ($M=282.5\%$, $SD=138\%$) proving again to be the most difficult tasks.

Pairwise comparisons (Bonferroni corrected) identify the *calendar* task to be the only one to differ significantly ($p=.014$) from the other tasks with respect to *percentage interaction steps*. This relative measure shows furthermore that the clear benefit of the gesture condition is no longer visible: not only is the most efficient interaction found in the gesture mode (170% of optimal interaction steps in the *message* task), but also the least efficient (299% in the *calendar* task). Comparing the two input modes for

each task separately, there is actually only a significant difference in the message condition where the gesture interaction ($M=170\%$) outperforms the key interaction ($M=253\%$), $z = -2.337$, $p = .019$, $r = -.41$; for the remaining tasks, no meaningful differences between the input modes were observed (Figure 7-11).

Figure 7-11. Percentage interaction steps.
Error bars denote a 95% confidence interval

Redundant screen changes. The measure of redundant screen changes completes the picture as again a significant main effect of task ($F(4, 124)=5.432$, $p<.001$, $\eta^2=.149$) and a significant interaction of task with the input mode ($F(3.1,96)=5.252$, $p=.002$, $\eta^2=.145$) can be observed. The number of unnecessary screen changes as a measure of efficient interaction was lower in the gesture mode than in the key mode for every single task (Figure 7-12).

Figure 7-12. Redundant screen changes.
Error bars denote a 95% confidence interval.

However, the differences between the two conditions vary considerably, ranging from 1.8 unnecessary screen changes (message task) to 7.8 additional screen changes (calendar task) in the keys condition. Wilcoxon signed-rank comparisons (exact, two-sided) for each task revealed that indeed only the median differences for the *calling* ($z = -3.231$, $p = .001$, $r = -.57$), *calendar* ($z = -4.732$, $p < .001$, $r = -.84$) and *website* ($z = -3.45$, $p = .001$, $r = -.61$) tasks can be assumed to be related to the different input modes.

Taken together, the analyses of the behavioral data in the logfile according to the five different tasks revealed that they were performed with varying efficiency. The first task, *calling a person from the contact list*, was performed fastest and with the least amount of unnecessary steps and screen changes. The tasks of *editing and sending a message*, as well as to *find and alter a calendar entry*, proved to be the most difficult. Across all dependent measures, a significant interaction between input mode and task was found, showing that a potential benefit of the gesture input identified in the previous section can be higher or weaker depending on the specific task. The difference between gesture and key input is usually largest for the tasks with which participants had difficulties in general, namely the *message* and *calendar* tasks. While the significantly faster completion of the message task with touch & gesture is probably due to the lack of excessive redundant key presses, the faster performance in the calendar task seems to be grounded in the huge difference in redundant screen changes.

7.3.3 Gesture Usage Patterns

Besides quantifying differences between the two input modes, it was also of interest to gain insights into the way *how* participants really interacted with the device. The following questions were addressed:

- How well were the gestures recognized by the system?
- Which gestures were used at all by the older users?
- Which gestures worked better, which did not work as well?
- Was there a difference between symbolic and direct manipulation gestures?
- Can the benefits of a touch interface be disentangled from the benefit of using gestures beyond simple taps?

Gesture recognition. Even though the gesture mode showed some advantages over the key mode in the efficiency measures and the user ratings, it is not yet clear how well the gesture control actually worked, in particular to which degree the gesture detection algorithm could reliably discern the implemented gestures. The algorithm compared each finger input pattern with a library of prerecorded patterns, and assigned a prediction score for each possible pattern, reflecting the probability that the current finger trace pattern is indeed one of the known gestures. If the winning

(=most probable) gesture had any functionality within the current screen, the effect was triggered.

Not in all situations was the appropriate or desired functionality triggered, which was either due to insufficient user input (e.g. an incorrectly drawn pattern), or due an insufficient detection performance of the software (e.g. the algorithm confused one gesture with another). The boundaries between these types of errors are fluent and hard to disentangle retrospectively, since there is no way to retrieve the actually *intended* gesture. Nevertheless, all 2124 recorded drawing patterns were visually inspected and manually classified as one of the fourteen possible gestures, as *artifact*, or as *unclear*, according to a set of specific rules (e.g., deviation from the horizontal by more than 45° disqualifies the pattern for swipe left/right). *Artifacts* are all those recorded patterns where no real movement trace was visible, e.g. when the screen was accidentally touched, or when an original tap, through minimal see-saw movements of the finger, triggered the gesture recognition system. *Unclear* gestures were all those patterns where a prolonged finger movement was recorded, but the human inspector could not unambiguously identify them to be one of the gestures from the implemented gesture set (see Section 7.2.1).

By comparing the classification of the human inspector and the classification of the gesture detection algorithm, we get an estimate of its performance. As *detection rate* we define the ratio of automatically recognized gestures of a particular kind to all gestures of that kind that, according to the human inspector, were drawn by the user. E.g. a detection rate of 50% means that of 100 patterns drawn in the shape of gesture X (as judged by visual inspection), only 50 were classified by the system as gesture X. Across all gestures, the system achieved a mean detection rate of 95.5% ($SD=5.8\%$). The gestures *swipe right*, *swipe down*, *arrowhead up*, *curve up* and *letter N* achieved perfect detection, whereas the lowest detection rate was calculated for the *question mark* gesture (80.5%). The individual detection rates can be seen in the confusion matrix (**Figure 7-13**), where drawn and detected gestures are contrasted and cell values mark the percentage of detection.

There is no one-dimensional explanation why certain gestures are erroneously identified as a different gesture. The confusion potential rises as the gestures become more complex, since sub-parts of the gesture might be classified as a different gesture. The most simple unidirectional swipe gestures show an almost perfect detection rate, while the more complex gestures like question mark, cross or heart are among the least reliably detected gestures. For example, an erroneous classification of an *arrowhead up* gesture instead of a *heart* gesture (6.3%) could have occurred if the two endings of the trace at the heart's apex did not close. Likewise, a *cross* gesture could be taken for a *curve up* gesture (13%) if the first stroke is performed bottom-up, and the delay before the second stroke gets too big. Not any pair of two specific gestures was

confused more than 13%. This analysis shows that even though the gesture detection in the prototype was not yet perfect and still leaves much space for improvement, it provided an acceptable level of reliability, with an average level of 95.5% correct detection over all gesture patterns.

		Detected gesture %													
		↑	↓	→	←	↷	↶	↗	↖	?	N	S	L	X	♥
Drawn gesture	↑	96,6				3,0		0,5							
	↓		100,0												
	→			100,0											
	←				99,2									0,8	
	↷					100,0									
	↶						95,6						2,2	2,2	
	↗						0,8	97,6							1,6
	↖								100,0						
	?		6,9				7,1	2,0		80,5			3,3	0,2	
	N										100,0				
	S								4,2			91,7		4,2	
	L												95,5	4,5	
	X					13,0								87,0	
	♥							6,3							93,7

Figure 7-13. Drawn gestures with corresponding detection rates (%).

Gestures used. The gestures were performed by the participants with greatly varying frequencies. The most abundantly used gesture was the *question mark*, which participants used on average 16.8 times, while the least used gesture was the *letter N*, which was used on average only by every 4th participant. These large differences are primarily due to the role of each gesture in the interaction and the task fulfillment. For example, the *question mark* gesture, available on every screen, was essential to learn about other gestures and how to interact with the system. The *letter N* gesture, on the contrary, had only one specific context (the contact list) and one specific purpose (to compose a message to this contact). In addition, its use was not essential for task completion, since alternative ways were possible. In **Table 7-2**, the estimated uses of a gesture and its actual uses are juxtaposed. To calculate the estimated number, necessary uses of a gesture were counted with 1, optional uses with 0.5. The values are averaged across the two different scenarios, which had slightly different demands. The actual usage value of a gesture is corrected by the detection rate of this gesture to obtain an estimate of the usage frequency assuming perfect detection.

Table 7-2: Frequency of gesture usage: *Est. uses* = estimated number of gesture use;
Avg. uses = averaged actual number of gesture use (perfect detection assumed);
Type = gesture type (*m*=manipulative, *s*=symbolic); *Ratio* = percentage of estimated gesture uses

Gesture name	Type	Est. uses	Avg. uses	Ratio	Remarks
swipe down	m	0,5	4,50	900 %	often tried out in order to scroll down
swipe right	m	1,0	7,53	753 %	often attempted to scroll right; correction for “overshooting” during side scrolling
curve down	m	0,5	1,34	269 %	attempted without centering on an object
arrowhead up	s	1,5	3,88	258 %	last resort if users got stuck
swipe left	m	3,5	7,72	221 %	redundant use in imprecise side scrolling
curve up	m	0,5	0,94	188 %	attempted without centering on an object
swipe up	m	3,5	6,13	175 %	often attempted on top of scroll bar
question mark	s	8,0	13,56	170 %	frequent usage to avoid memorizing gestures
letter S	s	1,0	1,38	138 %	sometimes repetition required because of scrolling effect
cross	s	1,0	1,25	125 %	additional attempts to close dialogs
heart	s	2,0	2,31	116 %	some repetitions necessary because of imprecise execution
letter L	s	1,0	0,66	66 %	contact list most often opened alternatively
letter N	s	0,5	0,25	50 %	contact usually selected directly from the message screen
arrowhead left	s	2,5	0,75	30 %	homescreen gesture preferred over back

Since the use of a particular gesture varies considerably between participants and is so intrinsically related to its specific use context and way of implementation, it is hard to derive any general conclusions. However, since the gestures were used on average more than twice as often as their estimated usage, the older users did not show a particular reluctance towards using them. Splitting the gestures in direct manipulation and symbolic gestures shows that an increased frequency of use is mainly caused by the direct manipulation gestures ($M=417\%$), whereas the symbolic gestures were approximately used within the expected frequency range ($M=119\%$). This difference could manifest the low thresholds these interaction patterns offer to explore the device: Participants often tried out whether a screen or list was scrollable, in which direction it was scrollable, whether an object could be moved etc. just by attempting a swipe or curve gesture. In comparison, the symbolic gestures had higher activation thresholds, partly due to their usually more complex pattern, partly to their more restricted use contexts. Especially the optional alphanumerical gestures were often bypassed by alternative interaction steps.

Noteworthy is furthermore the discrepancy between the *arrowhead up* (homescreen) and *arrowhead left* (back to last screen) gestures. Even though these two gestures are comparable in shape and effort, and have a similar functionality (returning to a previously encountered screen), the homescreen gesture was used much more often ($M=258\%$) than the back gesture ($M=30\%$), the latter even scoring considerably under the estimated frequency of use. It seems that participants, in situations where they felt stuck or lost, preferred rather to “go back to start”, trying the task anew from scratch, instead of backtracking to a previous screen from where alternative routes could have been chosen.

Recall rate. After the gesture interaction block and after the SUMI and QUESI questionnaires had been administered, participants were given a checklist where they had to determine for each gestures whether they remembered its functionality, and to rate the suitability of the proposed pattern for the assumed functionality on a 7-point Likert scale (not suitable at all – extremely suitable). Recall rates for a gesture’s functionality were very low: across all gestures, only every 4th user remembered a gesture’s functionality correctly ($M=25.6\%$). Disregarding the extremely popular question mark gesture, it is only every 5th user ($M=20.2\%$). The correct recall rate varies extremely between the different gestures, from over 90% for the question mark, to only 3.2% for the alphanumeric gestures (Table 7-3).

Table 7-3. Gesture functionality recall rates and perceived suitability ratings.
Gestures sorted by correct recall rate

	functionality recall (%)			suitability rating (1-7)	
	correct	false	unknown	<i>M</i>	<i>SD</i>
question mark	90,3	6,5	3,2	6,31	0,71
swipe right	48,4	22,6	29,0	5,82	0,59
cross	35,5	3,2	61,3	5,45	0,82
arrowhead up	32,3	3,2	64,5	5,00	0,95
swipe down	22,6	38,7	38,7	6,06	0,24
heart	22,6	29,0	48,4	5,56	1,15
swipe left	19,4	35,5	45,2	5,75	0,86
curve	19,4	0,0	80,6	5,00	1,29
swipe up	19,4	41,9	38,7	6,00	0,49
arrowhead left	12,9	16,1	71,0	4,64	1,36
letter S	3,2	25,8	71,0	5,10	1,29
letter L	3,2	3,2	93,5	5,75	0,96
letter N	3,2	0,0	96,8	5,33	2,08

The low correct recall rates have several reasons. Above all they are connected to the low frequency of use: some gestures were only used once or twice per participant, others not at all (cf. **Table 7-2**). Therefore, it is not surprising to see that for the majority of gestures (57.1%), the functionality was not remembered any more. Indeed, the rate of use frequency accounts for 69% of the variability in correct recall ($r=.83$, $p<.001$). A second reason is that participants were not instructed to remember gesture patterns or their connected functionality. During the experiment, there was also no need to remember, since participants could open a help screen any time during the interaction to obtain this information. The only gesture to be really remembered, therefore, was the question mark, which is reflected in its use frequency and recall rating.

Many gestures were not completely unknown, but associated with the wrong effect. The most prominent example by far was the direction of scrolling: by making an *upwards* movement, does the list scroll *downwards*? Or, in order to see the next pictures in the gallery to the *right*, is it necessary to perform a swipe *left*? This confusion was clearly visible during the interaction, and probably accounts for much of the observed increased use frequency for swipe gestures (**Table 7-2**). While during interaction through trial and error the participants learned how to navigate through screens and lists, they did not make this theoretical connection in the questionnaire and more often assumed, e.g., that a swipe up gesture would scroll up (42%) than scroll down (39%). Other sources of errors were that users were able to recall the context of use (e.g. "internet"), but not the specific functionality; that they confused the (visually rather similar) arrowhead up and arrowhead down gestures, and that they made up different associations for the alphanumerical gestures: two participants recalled the *letter S* gesture to have a *search* functionality, while the correct effect was *sending* an object.

Suitability. The suitability rating of gestures had to be evaluated on a 7-point Likert scale (1=not suitable at all – 7=extremely suitable; verbal anchors chosen after Rohrmann (1978) to be approximately equidistant). Almost all scores range between 5 (rather suitable) and 6 (very suitable), and no large differentiation can be seen (**Table 7-3**). The gesture judged most suitable was again the *question mark* ($M=6.31$, $SD=0.71$), followed by the swipe gestures ($M=5.75-6.06$). At the lower end we find the *arrowhead* gestures, indicating that these symbols might not have been as intuitive as the more familiar symbols like *question mark*, *heart* or *cross*. The suitability rating also shows a strong correlation with the frequency of use ($r=.71$, $p<.01$), showing that the suitability rating was not a static, but instead with increased usage, people started to learn the gesture's effect and consequently rated it more suitable.

Benefit of basic touch versus benefit of gestures. One of the novelties as compared with touch & gesture interfaces already on the market was the substantial use of

symbolic gestures, which had been (partially) derived from previous studies. It would be of interest to see whether the benefit of the gesture interface that has become apparent in the behavioral data and the questionnaire data (see Section 7.3.1) can be broken down to a benefit that was simply added through the basic touch interface, a benefit that could be attributed to the direct manipulation gestures, and a benefit which might have been added through the use of symbolic gestures. Because of the setup of the study and the measurement, their relative impact can only be estimated, rather than measured, based on their proportion in the entire interaction.

Each interaction step can be assigned to one of four basic categories:

1. **keystroke**: the participant pressed one of the physical keys on the device.
2. **tap**: the participant touched the display at one point with one finger (including long press)
3. **direct manipulation gesture**: the participant performed a finger movement on the touchscreen surface in order to manipulate objects or screens, e.g. swipe and curve gestures.
4. **symbolic gesture**: the participant drew an iconic or symbolic sign pattern on the touchscreen, e.g. a letter, heart, cross etc.

Keystrokes were not allowed in the gesture operating condition. They did occur sometimes by accident, but are negligible (0.13% of total interaction steps). The majority of interaction steps in the gesture condition were simple taps (52.5%), followed by direct manipulation gestures (26.1%) and symbolic gestures (21.2%). These numbers illustrate that gestures made up a substantial part of interaction steps (almost 50%) in the gesture condition, roughly equally divided over direct manipulation and symbolic gestures.

Figure 7-12. Frequencies of taps and gestures for the different tasks

Breaking up this distribution further according to tasks and gestures (**Figure 7-12** and **Table 7-4**), reveals that the relative use of gestures varied between tasks (35.7% gestures in the message task compared to 59.9% in the gallery task), and that the question mark gesture alone accounts for 45.5% of all symbolic gestures used. However, there was no significant correlation between the number of interaction steps and the touch/gesture ratio or the direct manipulation/symbolic gestures ratio.

Summing up, even though the proportion of gestures used in the interaction varies between tasks, there is no indication of a systematic change in interaction behavior when tasks get longer or more complex. Although the relative impact of touch vs. gesture interaction on global measurements of efficiency cannot be totally disentangled, both were used in approximately equal shares.

Table 7-4. Percentage of total interaction steps according to task

	taps	keystrokes	gestures			total gestures
			direct manipulation	symbolic gestures	question mark	
calling	60,6 %	0,2 %	16,5 %	22,6 %	11,1 %	39,1 %
message	64,1 %	0,2 %	17,7 %	18,0 %	6,6 %	35,7 %
calendar	58,0 %	0,1 %	25,2 %	16,6 %	9,3 %	41,9 %
gallery	40,1 %	0,0 %	42,4 %	17,5 %	8,6 %	59,9 %
website	51,0 %	0,2 %	13,0 %	35,8 %	14,8 %	48,8 %

7.3.4 Influence of Previous Knowledge and Experience

The participants had different levels of prior experience with computers, mobile phones, and touch & gesture interfaces (see Section 7.2.6). In this context it is of interest to assess whether the observed benefits of a gesture interface are likely to apply to a broad spectrum of older users with various levels of previous experiences with technical devices, or whether the measured benefits are rather due to a technology-savvy subgroup that could really take advantage of the gesture interface. To answer this question, the contribution of experience-related parameters, such as computer literacy, mobile phone experience, previous exposure to touchscreens or gesture interfaces, to the performance-related measures, such as time on task, redundant interaction steps, redundant screen changes, etc. was determined

The following measures were taken to assess previous experience:

- **Mobile phone experience:** Participants rated eleven typical tasks on a mobile phone according to usage frequency (0=never, 1=rarely, 2=occasionally,

3=often). The compound score is the sum of all frequencies, ranging from 0 (not a single functionality has ever been used) to 33 (all listed functionalities are being used often). This score can be criticized in two ways: First, it can be argued that the ordinal frequency data is not equidistant and thus cannot be used for addition. Arguments comparing the pros and cons of such a procedure and in support of the view that under certain conditions ordinal data can be treated like metric data are not discussed here, but can be found in Labovitz (1970) and Rohrmann (1978). Second, a non-weighted sum implies equal importance of the different tasks, which is surely not the case. For example, making use of the *calling* or *texting* functionality should be more indicative of general mobile phone usage than using the *games* functionality. Nevertheless, for lack of a theoretically grounded weighting procedure, this parameter still has its value in providing a quantifiable estimate of the participants' level of expertise with mobile phones: low scores indicate that participants use only a limited set of the possible features with low frequency, while high scores show that a broader range of functions is used more frequently.

- **Computer experience:** Computer experience was assessed through a subscore from the Computer Literacy Scale (Sengpiel & Dittberner, 2008), which is the sum score of correct answers from a test assessing knowledge of computer-related symbols and terms. This score is believed to reflect, rather broadly, how well people are experienced with computers and related technologies.
- **Touchscreen experience:** From a range of eight typical touchscreen devices participants indicated those they had already used, resulting in a value between 0 (no device used) and 8 (all devices used). This is a crude measure indicating the degree to which participants are familiar with the concept of a touchscreen, assuming that with a larger spectrum of devices used, the general understanding how such screens react and are to be used, is increased as well.
- **Gesture experience:** Only three out of 32 participants indicated to have used a gesture-controlled device before, and only one of them added a believable use case. Through lack of variance and validity, this item is ruled out as a good indicator of previous gesture experience. Instead, the users' familiarity with the interaction concept of the iPhone yielded a more fine-grained differentiation (see Section 7.2.6) to which degree they were familiar with the concept of using finger movements to control a mobile device. Gesture experience, in this context, therefore also includes having seen or heard how the product is used, not only having used it personally. This score is ordinal data, ranging from a minimum of 1 (=don't know the product) to 6 (=own the product).

A first step is to look at the correlations between the previous experience-related parameters and the gesture-related parameters described in Section 7.2.3. Moderate

negative correlations were found primarily between the computer literacy score, and the performance measures in the gesture mode, such as total task time ($r = -.44, p < .05$), redundant steps ($r = -.46, p < .01$) and step percentage ($r = -.46, p < .01$). This relationship provides a first hint that more technology savvy users (with higher CLS scores) could profit more from the gesture mode, since for the key mode, there were no correlations. Furthermore, a correlation between gesture experience and the total task time ($r_s = -.37, p < .05$) parameter was found, again only for the gesture condition, pointing in a similar direction.

Pursuing this idea further, a repeated measures MANCOVA was conducted, including the four performance measures described in Section 7.2.3 as dependent variables, the input mode as the independent variable, and the experience parameters as covariates. Following the arguments put forward by Thomas and colleagues (2009), we adopted the method proposed by Delaney and Maxwell (1981) to mean-center the covariates, thus getting rid of a main-effect distortion due to the repeated measures design (For discussion of the Delaney-Maxwell method, see Aligned, 1982). Besides the already discussed main effect of input mode (see Section 7.3.1) on the different efficiency measures (using Pillai's trace, $V = 0.86, F(4,24) = 37.12, p < .001, \eta^2 = .86$), a significant effect of mobile phone experience on efficiency measures was observed ($V = 0.36, F(4,24) = 3.371, p = .025, \eta^2 = .36$).

More interesting was the question if previous experience affected the interaction performance with gesture or with keys differently, and in particular, whether the efficiency advantages of the gesture condition were generated by the more experienced users. This hypothesis could not be confirmed since there were no significant interaction effects ($p > .05$) between the covariates and the input type. The previously reported correlations were reflected in the interaction terms of subsequent univariate ANCOVAs, such as an interaction of computer experience and input type on the percentage redundant interaction steps, and an interaction between gesture experience and input type on the total task time. However, if we Bonferroni-adjust the alpha level to control the family-wise error in multiple comparisons, these interactions do not reach statistical significance.

7.4 Discussion Study 3

7.4.1 Advantages of Gesture Control over Key Control

Summing up the results from the behavioral data and the subjective ratings, a clear indication in favor of the touch and gesture input can be observed. The behavioral

usability measures confirm that older users are faster ($p < .01$), use fewer unnecessary interaction steps ($p < .001$), use fewer relative interaction steps (*n.s.*) and perform also fewer navigation errors ($p < .001$) in the gesture condition than in the key condition. In addition, the SUMI and QUESI scores speak a clear language: In *all* of the 10 tested subscales, as well as the two global scales, the touch and gesture condition was rated better by the older participants (10 of the 12 comparisons were significant at $p < .05$). Finally, 72% of the older participants stated an explicit preference for the gesture control over the key control. This strong preference of gestures over keys cannot be taken for granted, since 94% of the participants possessed a key-controlled mobile phone, only 17.1% stated they were familiar with touch-controlled mobile phones, and 84.4% stated they had no idea of the basic interaction mechanisms of an iPhone.

Not surprisingly, the overall advantage of the gesture control did not arise equally from the different tasks: some tasks lent themselves better to gesture input, others less. Still, four out of five tasks were performed faster, and all tasks showed fewer redundant interaction steps and fewer navigation errors (redundant screen changes) in the gesture mode. The benefit of the gesture interaction became particularly clear for those tasks which were difficult in general, in our case the message and calendar tasks.

On a first note, the results are very promising regarding the senior-friendliness of gesture interaction. It seems to be the case that this type of interaction could indeed facilitate technology interaction for older adults, at least in the case of mobile devices, and in comparison with the classical way of interfacing such devices by means of keys and menus. However, before any clear-cut design recommendation can be given, alternative explanations need to be ruled out and the limitations of this study have to be discussed.

7.4.2 Learning effect

Since participants performed similar tasks in both conditions, order effects such as carryover, practice or fatigue should be controlled for. In order to minimize the influence of these effects, the order of input modes was counterbalanced between participants. Since there was no significant interaction of order and input type on any of the behavioral dependent measures, we can assume that the manipulation was effective and effects of input type can be interpreted taking both orders together.

However, the measure of redundant screen changes showed a main effect of order. If keys were the first interaction and then gesture, participants made fewer errors in the second gesture condition compared with gestures as first condition. This can still be interpreted as an effect of practice, since participants had already gained some experience with the device through prior key interaction. However, errors did not decline

in the key condition if this condition was presented second. On the contrary, screen change errors occurred more frequently as compared to the condition where key interaction came first. This suggests that prior gesture interaction experience is detrimental to subsequent key interaction, when looking at the overall smoothness of navigation. Taken together, no overall practice effect could be substantiated over the two repetitions of task. Since both interaction conditions were so fundamentally different, it is not surprising to find little carryover and learning between the two repetitions, despite the highly similar tasks and the same screen layout.

Yet, a very comparable study (Schnock, 2011) using the same functional prototype, but *four* task repetitions and only the gesture interaction paradigm, could substantiate a significant improvement in *time on task* (58%), a significant reduction in *interaction steps* (33.7%-43.9%), and a significant drop (56%) in the *rate participants were calling the help function*. Also the subjective ratings on the *mental demand*, *performance* and *frustration* subscales from the NASA TLX questionnaire (Hart, 2006; Hart & Staveland, 1988) underline the learning effect: frustration decreases, mental demand declines significantly, and performance increases significantly over the four task repetitions (Schnock, 2011). Therefore, even if the two repetitions in this study were too few to show measurable and systematic learning, and the effects of order were dwarfed by the effects of input type, as few as four repetitions suffice to demonstrate performance increases of more than 50% for older users. This result is encouraging, since it shows that not only is gesture interaction beneficial for a novice older adult who uses such a device for the first time, but it also permits a steep learning curve, where older users become dramatically faster and require less and less help after only a very limited number of interaction episodes. The ease of learning has been demonstrated to be an important predictor of acceptance or rejection of technology among older users (Renaud & van Biljon, 2008). This being said, a direct comparison of the learning curves in gesture control and key control for older adults is still lacking and subject to future research.

7.4.3 Limitations of the Approach

The results of study 3 clearly show that the older participants favored one of two tested interaction paradigms over the other. However, the extent to which gesture interaction alone can account for the observed differences, and the extent to which the findings can be generalized to other devices using gestural interfaces, requires a reflection.

Single finger gestures. The functional prototype used a touchscreen capable of sensing only a single finger at a time. Therefore, the potential that multitouch offers with respect to direct interaction – e.g. rotating, repositioning and zooming objects

seamlessly with 2-finger gestures – could not be fathomed for older users. In situation where single-finger gestures interfere with scrolling, multi-finger gestures could sometimes be viable solution, which also was not an option in the current prototype. In such cases, other workarounds had to be found for this study. Consequently, the gestures that were tested in this study are to some extent the result of technical limitations, and cannot claim to be the best possible gesture set. Even though two-finger gestures, if tested in isolation, were less preferred by older adults (cf. Section 6.9.1), a gesture set containing also some two-finger gestures might still be superior considering the interplay of gestures within the whole set.

Gesture recognition. The gesture detection mechanism used in the functional prototype was not flawless. It was based on the algorithms provided within the *android.gesture* libraries. The gesture recognition algorithm re-samples the input pattern, geometrically transforms it (e.g. centering, rotation and scaling) and matches it to patterns stored in a previously established library, based on similarity measures like Euclidean and Cosine distance between two gesture patterns. Based on the similarity measure, each gesture in the library was assigned a prediction score, and the gesture with the highest prediction score was selected and executed by the system. Three disadvantages were coupled to this system.

First, since only the name or ID number of the best matching gesture was available from the system, gestures were working much like commands: if a certain gesture was detected as best prediction within a certain context, a specified effect was triggered (e.g. to switch to another screen). However, a real-time direct manipulation of objects, e.g. the simultaneous dragging of a screen object along the fingertip, was not possible with this implementation.

Second, the gesture detection system lacked any adaptive elements. Users had to learn to adapt their patterns to the patterns previously stored in the gesture library, to which they had no access. A gesture detection system capable of learning – e.g. extracting common features from multiple instances of user generated gestures – would have much greater chances of reliably detecting the gestures of a particular user. Instead of a single template (or, a very limited number of templates) for all users, each user could enter a series of gestures within a short training session, of which the system generates the prototypical gestures for this specific user, against which all subsequent gestures are matched.

Third, gesture detection was suboptimal in the presented study. Even though the average gesture detection performance was estimated at 95.5%, which seems acceptable, the performance varied between 80.5% and 100% depending on the pattern. Especially the by far most frequently used gesture (*question mark*) was found to be detected worst (80.5%). A gesture that triggered the desired effect (pop up of the help screen) only in 4 out of 5 cases, and otherwise was not responsive, or worse, trig-

gered a completely unrelated function, caused sometimes confusion among participants. The high frequency of the question mark gesture could even be partly attributed to the fact that it was not recognized reliably.

Taken together, a more elaborate gesture detection mechanism, which cannot only handle static patterns but allows for dynamic content manipulation, which can be adapted to a particular user's style of gesturing, and which shows a consistently high recognition rate, will further improve the handling of a gesture controlled device, and thus also its probability of adoption by older adults.

Overlapping gesture patterns. Another technical limitation of the prototype and the way gesture detection was implemented is related to the scrolling/gesturing conflicts. Iconic and symbolic gestures cannot be trivially performed on screens that are also scrollable. For example, the first upstroke in a *Letter N* gesture could start vertical scrolling in a list before the pattern is completed. These conflicting situations were avoided in the prototype where possible, but therefore also the choice of appropriate gestures for certain contexts was limited. So far, no simple solution to this problem exists. In fact, the lack of symbolic gestures on currently commercially available gesture-controlled devices might be essentially grounded in this conflict.

One solution already discussed above is the introduction of multi-finger gestures. Another solution, implemented for example on Synaptics touchpads (Synaptics Inc., 2010), is to precede all symbolic gesture input by a specific gesture command (e.g. a three-finger tap), which signals the system that a subsequent finger movement should be interpreted as a gesture command. Other conceivable solutions could smartly consider the context in the gesture detection algorithms, so that for example a different gesture is executed depending on whether the gesture was started right on top of a list item or in between list items. Still, recent reports of gesture patents filed by Google hint in the direction that also global IT companies strive to include symbolic gestures (alphanumeric symbols, in this case) in their future interfaces (Purcher, 2012).

Within the scope of this work, some results (e.g. the over proportional use of the *Letter S* gesture) must be understood with these limitations in mind, and the resulting user confusion could have influenced gesture performance and ratings.

Interaction tree and screen design. The way the interaction tree (Figure 7-3) and screen design was chosen implies a critical account of the presented data. Alternative routes to arrive at a specific screen were explicitly implemented to ensure realism and accommodate alternative conceptual models which different users might have based on their prior experience. The drawback of this approach is that different users navigated with different screen successions through the interface, which renders the comparison of interaction steps slightly fuzzy. Furthermore, some routes contained

gestures which others did not; therefore it was not ensured that all users used all potentially available gestures. The frequency of gesture usage is thus not only an indication of how well a gesture was received, but also confounded with its obligatory or optional use.

The screen design is not free of bias for one of the two alternative input modes. Ideally, the input mode should have varied independently of the screen design, and each screen should not favor any of the two input modes. However, how well a system can be operated by a certain input paradigm is intrinsically related to the screen design and cannot be separated. Since screen design should not confound the comparison, it was held constant in both input modes to the greatest extent possible. Furthermore, care was taken that each screen can be operated reasonably well in both input conditions.

Nevertheless, some screens were undoubtedly easier to handle in the gesture mode. Particularly the tasks in which the strongest differences in behavioral measures were observed exemplify this. For example, the send message task contained a subtask of editing a message. While in the gesture condition the text editing could in principal be achieved with a *swipe* gesture (mark text passed), a *cross* gesture (delete passage), and a *tap* on the confirm button in the dialog box, the task in the key condition demanded explicit positioning of the curser using the navigation keys, and then cumbersome repeated pressing on the delete key. On the other hand, some tasks were easier using the keys. For example, the homescreen or previous screen functionality was always present and visible because of the dedicated hard keys, whereas nothing on the screen afforded the back (*arrowhead left*) and homescreen (*arrowhead up*) gestures, and they were easily forgotten and had to be looked up again.

Summing up, the degree to which gesture interaction turns out more or less favorable in comparison with key interaction is not only connected to an appropriate gesture set, but also to the general GUI design and the key layout. Some of the observed differences can be clearly attributed to implementation details of the prototype. Thus, evaluations with different screen designs and, more specifically, involving screen designs which are optimized for a certain input method, need to be carried out before the present results can be generalized to a broader range of applications and devices.

Key layout. Another limitation of the study was the fixed layout of keys. While the gestures and the screen design could be modified, the size, layout and design of the hardware keys was fixed. Particularly for older users, the usage of large keys with clear markings and appropriate interkey spacing is recommended (Fisk, et al., 2004). The keypad of the prototype (**Figure 7-4**) lacked interkey spacing, clear visual markings, clear tactile boundaries of some keys, and an unconventional ordering of the keys. Furthermore, the size of the keys, especially of the important navigation keys,

was too small (approximately 5x5mm). Participants reported difficulties to hit the right key, or accidentally slipped from one key to an adjacent key. Particularly often (and particularly detrimental) were slips from the right navigation key to the adjacent back key, which usually meant that participants found themselves on the previous screen (sometimes unaware why) and considerable progress towards goal achievement was lost. Such system behavior was of course frustrating, and can partly account for the increased time demands, increased number of interaction steps, and worse subjective ratings in the key input condition. The clear advantage in favor of the gesture interface that was documented above must therefore be put into perspective of the suboptimal key layout. Differences might not have turned out so clearly had the keypad provided sufficiently large and sufficiently spaced keys with a clear and high contrast labeling.

Touch and/or gesture. The study compared the interaction through keys with interaction through *touch & gesture*. In the context of this work, touches are considered the most basic and simple form of gesturing (the *tap* gesture), as argued for in Section 3.1.3. There is, however, a difference between touches and other gestures, since many devices exist that react to touch only. Examples are the touchscreens of ATMs, info terminals, or ticketing machines, but also of some mobile devices. Virtual buttons are their only interactive elements. In the present interface, virtual buttons (e.g. the OK and CANCEL buttons that appeared to confirm or reject an action) were mixed with direct manipulation gestures (e.g. scrolling) and symbolic gestures. The benefit of this mixed approach is a greater degree of realism, since most modern smartphones are operated through a mixture of virtual buttons and direct manipulation gestures. The disadvantage is that in our study we cannot disentangle separate contributions to the overall benefit of touch & gesture that was quantified. It might be the case that just the use of a touchscreen, on which elements can be selected directly without the cumbersome use of navigation keys, and where virtual buttons can be placed dynamically respecting the context and scaled with regard to user needs, already accounts for all the benefits in behavioral and subjective measures. However, it could also be the case that direct manipulation gestures such as swipes and curves add an additional benefit on top of the virtual button concept, or even that symbolic gestures add a benefit on top of the touches and manipulative gestures. But it could equally be that these gestures are actually detrimental and a pure touch-based interface would have turned out still superior. Further investigations in which these input types are separated are needed to quantify their relative impact on the overall increased usability found in the present study.

7.4.4 Usage Patterns

Preference of gestures types. The only indication of how well touches, manipulative and symbolic gestures were received in this study can be obtained by quantifying their frequencies of use in relation to their estimated frequencies. Taps, in the shortest interaction path scenario, have been estimated to be used on average in 59.6% of all interaction steps and other gestures in 40.4% of all interaction steps. The actually performed interaction steps showed a ratio of 52.5% of taps as opposed to 47.3% other gestures. This increased use of other gestures could either, in a positive reading, be understood as a special affinity for this type of interaction, or, in a negative reading, indicate that more repetitions were necessary because of erroneous detection, and an increased use of the question mark gesture because patterns were hardly remembered. Since this cannot be discerned, only the general statement can be made that older users did use gestures to a significant degree (roughly 50% of interaction events), and that it can be assumed that their subjective ratings and preference statements are also very likely based on their gesturing experience and not just on the touchscreen experience. Moreover, since direct manipulation gestures (26.1%) and symbolic gestures (21.2%) were roughly used to equal degrees, one can further assume that also both gesture types influenced the ratings.

Looking further into different types of gestures, we observed that direct manipulation gestures were used disproportionally more often (417% of the their estimated frequency) as compared with symbolic gestures (119%). This has been interpreted to reflect the lower activation threshold, making such gestures prime candidates for trial and error (e.g. to find out the appropriate scrolling direction), unlike the symbolic patterns (cf. Section 7.3.3). Another argument for the users' preference of direct manipulation gestures over symbolic gestures is the low usage frequency observed for optional symbolic gestures: they were bypassed and other navigational routes taken if these were available. Such a reading would be in line with the suitability ratings in study 2B, where older adults also rated indexical gestures (taps) and manipulative gesture better than symbolic or iconic gestures.

Suitability of individual gestures. Concerning the suitability of individual gestures, the *question mark* gesture used to trigger the help screen was by far received best. It was not only the most frequently used gestures, it was also the one which received the best suitability rating and the only one where its functionality was generally correctly recalled (90.3%).

Swipe gestures were the second most frequently used gestures. Interestingly, confirming the observations of study 2A and 2B, even within the functional prototype the scrolling direction was problematic to the older users. In some tasks, goal relevant items were hidden at lists positions which required the user to scroll down from

the top of the list with the help of a *swipe up* gesture. Still, *swipe down* was found to be the gesture with the highest disproportion of minimal (0.5) and actual (4.5) number of uses. The misconception of scrolling direction can also be derived from the fact that *swipe up* and *swipe left* gestures (scroll down and scroll right) were required to be performed much more frequently, and still *swipe up* and *swipe right* received the higher usage ratios (see **Table 7-2**) and higher suitability ratings (see **Table 7-3**).

The *cross* gesture ranked third in the list of correct recall (35.5%) and also received average suitability ratings. Since it had to be employed only once in the usage scenario, further evaluation is needed, but prospects are positive that it could be well received.

Similarly, the *heart* gesture, which had to be performed at least twice, showed a comparably well correct functionality recall (22.6%) and received the best rating after the question mark and *swipe* gestures. Considering the fact that the pattern was slightly more complex than most other gestures (two strokes, including curvature, symmetry and closure) this gesture was well received, but further testing is needed.

Interesting to observe was the clear preference of the home gesture over the back gesture, despite the similarity of the pattern (*arrowhead up* vs. *arrowhead left*). The back gesture was used less than half as often as expected, while the home gesture was used more than twice as frequently as expected. It seems that older users, in situations where they feel stuck or lost, prefer to “go back to start”, trying the task anew from scratch, instead of backtracking to previous screens from where alternative routes could be chosen. Quite possibly it was also less clear to the participants on which screen they will land if they perform the *back* gesture, whereas they knew exactly what would happen upon performing the *home* gesture.

At the lower end of the spectrum we find the alphanumeric gestures *S* (“*S* end”), *L* (“*L* ist”), and *N* (“*N* achricht”, i.e. “message”). They were used least frequently (because of alternative routes of navigation), they were hardly recalled (3.2% correct), and received below-average ratings. One reason for this result could be found in the inability of the participants to make the association between the gesture letter and the connected command. Even though the connection was highlighted in the help screen, it was neither explained, nor trained and repeated to a sufficient degree for the older users to internalize the association. The alphanumeric gesture commands might thus provide shortcuts for a trained and experienced user, but are less intuitive for the novice older user. Only after intensive trainings and repetition might they show any benefit, if at all. This observation is relevant for future gesture interfaces which rely stronger on alphanumeric input, such as the smart search function suggested by Google (Purcher, 2012).

7.4.5 Influence of Prior Experience

In order to make inferences to which degree the observed results can be generalized to an older population with varying degrees of technology experiences, the relative impact of prior experience with computers, mobile phones, touchscreens in general and gesture interaction in particular, was assessed. From correlation analyses some indications were obtained that particularly computer literacy was related to a better performance in the gesture mode. However, no such systematic relations were found for other experience related parameters, and also, no such relations could be substantiated for the interaction performance with keys. MANCOVA analysis confirmed an influence of prior experience with mobile phones on the efficiency measures, but no interactions between prior experience and the input type were found. Taken together, some indications of an influence of prior experience with technology were obtained, but these were neither systematic, nor found to affect key and gesture performance dramatically different. In particular, an increased knowledge of touchscreens and gesture interaction did not result in a unilateral performance increase in the touch & gesture input condition. A restricted recommendation of touch & gesture technology only for technology-experienced users seems not necessary. On the contrary, the relative independence of performance from prior experience recommends such devices also to novice older users with little prior experience with related technologies.

7.5 Conclusions : Touch & Gesture outperforms Key Interaction

This chapter comprised the implementation and evaluation stages of the project as outlined in Section 4.3.2. A functional prototype of a gesture-enabled mobile phone was developed based on an Android smartphone with touchscreen and hard keys. A gesture set which was partly derived from the findings of study 2B was implemented and compared with conventional input through keys. While studies 1A-2B set the ground for a prospective positive user experience of the older adults regarding gesture interaction, study 3 was the first in which gestures had to be performed in context and where they actually triggered an effect. As comparison served the interaction through navigation and function keys, since this is still standard on many modern mobile devices.

Despite many limitations of the prototype (see Section 7.4.3), the results speak in favor of the gesture interaction paradigm. Effectiveness was not really measured, but perceived goal achievement was significantly higher in the gesture mode. Errors, quantified as unnecessary deviations from the optimal navigation path, were signifi-

cantly increased in the key condition, and so was the perceived error rate. The older participants were not only significantly faster in the gesture condition, they also needed considerably fewer interaction steps to fulfill the given tasks, and subjectively rated the efficiency of the system higher. The older users also rated the interaction with gestures more intuitive and better learnable than the key control, even though they had far more previous experience with key controlled mobile phones. An overall acceptance of the system was hard to assess after only a single interaction episode. Nevertheless, the older users found the gesture interaction significantly more helpful, more likeable and more controllable. Finally, 72% uttered a preference for the gesture system, compared with 22% preference for the key control and 6% indifferent. The majority (59.3%) agreed or completely agreed with using such a gesture controlled device in the future.

The analysis of different gestures proved a preference for easy manipulative gestures. However, simple iconic and symbolic patterns with clear associated semantics, such as *question mark*, *heart*, or *cross*, were also often used and favorably rated. In particular, using the question mark gesture to bring forth a context-aware help screen at any point in the system was very well received and can be recommended to be implemented in future systems in a similar way. The other gestures need still further evaluation since they occurred only rarely in this study. Alphanumeric gestures are in general not recommended to be used in gesture-control systems for older adults, unless an intensive training can be provided to learn the match of gesture patterns and effect. Providing an easily accessible gesture to return to the home screen was found to be essential for older users, while the possibility to backtrack to a previous screen by means of a gesture was rather disregarded, despite the use of a gesture similar to that used to return to the home screen. In line with the previous studies, the results of study 3 underline the confusion among older users regarding the appropriate scrolling direction. Here again, strong visual aids and sufficient training is advised for older adults to shift from their prevailing windows conceptual model to the paper metaphor of scrolling.

In conclusion, the study has demonstrated that the usability of a touch & gesture controlled device can live up to, and even surpass the usability of a key controlled mobile device for healthy older adults. As discussed in Section 7.4.3, many design particularities of this study could account for the observed effects, and further evaluation is thus necessary before the results can be generalized and cast into universal design recommendations. Still, the foundation is laid to include gesture interaction in senior specific devices with the prospect of facilitating interaction, and making technology more appropriate to support older adults in their daily routines.

8 General Discussion and Outlook

At least two global tendencies give a strong motivation to investigate the way how older people interact with technology, and how their interaction can be facilitated. First, in western societies the number of senior citizens is rapidly growing, thus older adults will also become increasingly relevant as potential users of technology. Second, technology expands into more and more areas of our daily life, for example through smart homes, ubiquitous computing, mobile computing, social networks, location-based services, to name but a few recent developments. The confluence of these two trends bears great chances for an aging population, but also great risks. If new technologies do not bypass older adults, but instead are designed to match their needs and abilities, they could support elderly in various daily routines, e.g. in communication, information, navigation, healthcare or safety, and foster independent living. On the other hand, if older adults are cut off from recent technological advances, they may run the risk of being dissected from society as well, e.g. in terms of information or communication channels, and the so-called “digital divide” may widen (Czaja & Hiltz, 2005).

Fathoming the potential of gesture interaction to bridge this divide was the driving force behind the present research project. The first three sections of this chapter summarize the empirical results, the limitations of the research and the main contributions. The last two sections address the practical implications for inclusive gesture control and future directions of research.

8.1 Summary of Results

The main findings from the five empirical studies are briefly summarized, before differences and similarities across the studies are presented.

8.1.1 Stage 1: Age Effects on Gesture Execution Performance:

- Screen size influences gesture performance across age groups. Smaller screens provoke more erroneous gestures
- Screen size influences gesture speed. The larger the screen, the faster the gesture is executed

- Screen size influences gesture accuracy. The larger the gesturing space, the more accurately gesture patterns are reproduced
- The difficulty of a gesture pattern is determined by more than the number of lines of which it consists. A multi-dimensional account of gesture complexity is needed that takes into account parameters such as curvature, symmetry, and direction
- More difficult gestures are performed slower and less accurately
- The way a device is held and operated influences gesture performance. Single-handed operation posture (gestures have to be performed with thumb) is particularly detrimental concerning accuracy, errors and speed
- One-finger gestures can be performed reasonably well, if the device is held in one hand and the other hand performs the gesture. Two-finger gestures are more erroneous, slower, and less accurate in handheld operation posture compared with a stabilized posture
- Device operation posture affects both age groups in the same way
- For two-finger gestures, a stabilized posture is preferred across age groups
- Older users perform gestures slower than younger users, but the slowing factor of 1.3 is low in comparison with the slowing factors of other input devices. Younger users are particularly faster on larger screens
- No systematic or substantial age group differences regarding the accuracy of gesture performance could be measured
- Only two-finger gestures are performed less accurately by older users
- Older users perform more erroneous gestures. Problems seem to be related to memory, attention span and fatigue rather than insufficient motor performance
- The subjective difficulty of a gesture pattern is perceived similarly in both age groups
- Upward and leftward gestures are subjectively perceived as more difficult than downward and rightward gestures
- Older users report curved gestures as particularly difficult
- Older users rate iconic and symbolic gestures better than younger users

The findings show that several factors influence the performance of finger gestures on a touch sensitive surface. The central question of interest, whether gesture interaction has such demands on manual dexterity and motor control that older adults face substantial problems to properly execute the gestures, needs to be negated on the base of the present data. Older adults perform gestures slower, but no evidence of a systematic drop in accuracy – which is most important for automatic gesture detection – could be found. Even with severely limited screen space or restricted thumb operation, no interactions with age group were found. The gesture syntax (at least to

the extent of simple geometric patterns like the ones tested here), therefore, does not need to be drastically modified to build a set of gestures tailored specifically to include the elderly. In fact, from the point of view of accurate pattern reproduction, the same gesture set can be used for older and younger user groups. However, if design recommendations should be derived, two-finger gestures should be avoided, curved gestures sparsely used, and the application of iconic and symbolic gestures reconsidered for a senior inclusive device.

8.1.2 Stage 2: Intuitive Gestures for Older Users

General findings:

- Little agreement exists on the most suitable gesture for a particular task across age groups
- Indexical and manipulative gestures, which are the least complex, are most frequently suggested, and also generally rated more suitable than symbolic and iconic gestures. They are predominantly selected as best fitting gestures, especially the most parsimonious tap and swipe gestures
- Gesture creation and gesture evaluation result in different gesture sets: while excogitating gestures has the tendency to result in extreme diversification, evaluating a predefined selection has the tendency to over-simplify different approaches and to sacrifice specificity for the sake of simplicity
- Gestures should be evaluated in combination instead of individually

Age-related findings:

- Different gesture types are suggested by older and younger adults. Older adults suggest more iconic and alphanumeric gestures, while younger adults suggest more manipulative gestures
- Older and younger users differ in the type of gesture they prefer
- Younger adults prefer two-finger gestures over one-finger gestures, while older adults prefer one-finger gestures.
- Older adults suggest more complex gestures, and also perceive complex gestures more suitable than younger users
- Older adults show less agreement than younger adults concerning the appropriateness of a gesture
- Familiarity of a gesture patterns is more important for the older users than for younger users
- Age as well as prior experience with related technology correlate highly with gesturing style (e.g. types used, complexity, number of fingers used)

- Prior gesture experience seems to determine the preference for two-finger gestures, while the preference for gesture type (manipulative, symbolic ...) is rather age-related
- Prior gesturing experience rather than age seems to account for the fact that older users favor a different scrolling direction (windows metaphor) than younger users (paper metaphor)

A set of intuitive gestures is hard to establish based solely on user input, as little agreement exists on the most suitable gesture for a particular task. Although direct user input in determining a set of suitable gestures is certainly important, the results are to be used as a source of inspiration rather than a directive for system designers.

Taken together, a gesture interface which relies strongly on indexical and manipulative gestures can be recommended for both age groups, and no totally different gesture set must be developed for older users. However, if in doubt, a system designer can trade an arbitrary multi-finger pattern for high familiarity, strong iconicity or a well known symbol, even if the pattern complexity increases slightly.

8.1.3 Stage 3: Evaluation: Gesture Control Outperforms Key Control

- Older users are faster with gesture interaction compared with key interaction on a mobile device
- Older users are more efficient and make fewer errors with gesture interaction compared with key interaction
- Older adults subjectively attribute a higher usability and intuitiveness to gesture interaction as compared with key interaction
- Older adults perceive gesture interaction as more efficient, more helpful, more likeable, more controllable, easier to learn, more familiar and less error prone than key interaction
- Efficiency and errors of gesture and key interaction vary substantially across different tasks
- Gesture interaction can be learned quickly by older adults. As few as four repetitions suffice to reduce time on a task, interaction steps, and need for help by 40-60%
- Older adults prefer touch & gesture interaction to key interaction
- Older adults prefer manipulation gestures to symbolic and iconic gestures
- Older adults prefer returning to the home screen to returning to the previous screen
- The question mark gesture for context sensitive help is well received by older adults

- Older users perceive alphanumeric gestures as little suitable

In direct comparison, gesture interaction on a mobile device was performed and perceived better than conventional key interaction. Even though many particularities of the tested prototype can account for the preference of touch & gesture, and generalizations to different devices and gestures require further empirical evidence, a carefully positive conclusion can still be drawn: The novel input paradigm to operate devices by means of finger gestures on a touch-sensitive screen seems at least not to place a new obstacle between the elderly and technology. On the contrary, it seems to facilitate interaction, and is furthermore positively received among older adults.

In all five empirical studies, participants were asked, whether they would like to use such a system in the future, and in all studies, the majority of older participants indicated their willingness to use such a system (**Figure 8-1**).

Figure 8-1. Willingness of older adults to use devices with finger gestures.
 4-point Likert scales (studies 1A & 1B, no "neutral" option);
 5-point Likert scales (studies 2A, 2B & 3);

In studies 1A & 1B usage intention was assessed on a 4-point Likert scale (strongly agree - strongly disagree). The initial elimination of the neutral category was intended to force the users to take sides and obtain a clearer picture about the user's opinion on gesture interaction. The surprisingly high level of agreement (some research argues that the midpoint omission could increase respondents bias towards positive replies, Worcester & Burns, 1975), and the comments from participants that they really had not enough information to judge after only one brief exploration phase (Tull and Hawking (1993) argue that an omission of the neutral category will provide less accurate responses if the respondents truly have no opinion on a subject) led to the inclusion of the neutral category for studies 2A, 2B and 3. The results

visualized in **Figure 8-1** are surely distorted to some degree by acquiescence bias and social desirability bias, but the fact that in all five studies the majority of older adults stated that they would like to use such gesture-controlled systems in the future, can be seen as an indication that they perceive a benefit and are willing to accept rather than reject such device.

8.2 Limitations of the Research Approach

Taken together, the results gained from the empirical studies are encouraging to use touch and gesture input also, and specifically, for devices dedicated to older users. No particular difficulties in drawing the gesture patterns could be substantiated, no major restrictions to a general gesture set had to be made, interaction was measured to be comparably usable, and was simultaneously positively received. However it would be too early to formulate unconfined recommendations for implementing gesture technology in a broad range of devices and for a broad range of older adults. The sample of participants was limited in many ways, the gestures tested were very specific, and general limits of cross-sectional studies and laboratory settings apply. For a better estimate of the momentousness of the presented data and its implications, some general limitations of this research approach are discussed in the following.

8.2.1 Sample Structure

Sample size. Studies 1A and 1B resulted in largely non-significant effects. Sample sizes ($N=36$ in 1A, $N=39$ in 1B) had been planned to reveal medium main and interaction effects ($\eta^2 = 0.0588$, Cohen, 1988) for the $2 \times 3 \times 3$ repeated measures design at $\alpha = .05$ and $1 - \beta = .08$ (cf. Bortz & Döring, 2006). However, in both studies, most of the age group effects and interaction effects with age turned out to be small (**Appendix A.2** and **A.3**). Given the generally high correlations between the repeated measures, the tests should still have enough power to reveal also smaller effects ($.02 < \eta^2$), yet to reliably determine all main effects of age group, larger sample sizes would have been needed.

Gender. Across all studies, 116 older adults were analyzed, which were gender-balanced. Gender differences regarding for example computer anxiety, computer self-efficacy or interest in computers, as well as interactions with age (computer anxiety) have been demonstrated previously (Czaja, et al., 2006). However, the present work followed an applied perspective, according to which gesture-controlled devices should be usable by both genders alike, and therefore any gender specificity

was not investigated further. Instead, care was taken that in all studies the age groups were gender-balanced, and between-subject manipulations were counterbalanced across gender to avoid distortions of the data.

Physical Ability. The age range of the older participants was 60-75, with an average age of 66.6 years ($SD=3.9$ years). Older participants were generally in good health conditions: 98.2% reported normal or corrected to normal vision, and 91.4% reported no restrictions in manual dexterity, hand and finger movement or wrist flexion. Those participants who stated minor restrictions reported that this did not affect the experimental task demands. The vast majority were right handed (97.4%).

Education level. Previous research has already documented the impact of education on computer knowledge and its strong mediating role for computer anxiety, as well as the negative relation between age and computer knowledge (Ellis & Allaire, 1999). Czaja and colleagues (2006) also found the positive impact of education (besides the negative impact of age) on technology use of older adults.

The sample of older adults was exceptionally well educated: While in our sample almost 47.4% had a university or comparable degree, the average in Germany is only 26% among the 55-64 old (and only 24% in the OECD average, Statistisches Bundesamt, 2011). The disproportionate fraction of participants with tertiary education can be explained partly by the way of recruitment through a newspaper advertisement and an online test person database. The database had been filled primarily with visitors of a science fair and required online registration. The newspaper advertisement offered participation in university studies concerning older adults and technology. Thus, participants are likely to have relatively high levels of affinity for science, university studies, and on top of that, affinity for technology. All of this is reflected in their relatively high education levels.

Thus, it is probable that a different sample, mirroring the distribution of education levels in the population, would have shown a less favorable interactive behavior and subjective reception of the novel input paradigm. As a result, caution applies to all generalization concerning the appropriateness of gesture input to all elderly in this age band. On the positive side, also the young participants were disproportionately highly educated (74.7% students), which means the cross-sectional comparisons (1A-2B) should not be strongly confounded by education.

Motivation. High motivation of participants could have been a distorting factor in favor of a positive reception of gesture technology. Motivation levels were not systematically assessed or accounted for. However, the fact that the participants were volunteers and their interest in technology (see previous section) may have caused above average motivation levels. Although motivation varied, most older participants uttered a strong interest to learn something about novel technologies in general

or cell phone and smartphone usage in particular. A few older adults explicitly stated they were considering buying a new phone and wanted to seize the chance of a “guided training” before making the purchase decision. Melenhorst, Rogers and Caylor (2001) have shown that such *perceived benefits* are a motivational factor for using or not using a technology, and especially so for older adults who base their choices often on a *cost-benefit-analysis* (Melenhorst, et al., 2006). Besides the small financial reward, participants thus can be assumed to have had a strong intrinsic and extrinsic motivation. High motivation could, in parts, account for the high accuracy of gesture performance (studies 1A-1B), the large number of creative gesture suggestions (study 2A), and the overly positive evaluation of gestures (studies 2B & 3).

Since all of these factors have been demonstrated to foster the intention to use information technology (Nysveen, Pedersen, & Thorbjørnsen, 2005; Renaud & van Biljon, 2008; Venkatesh, et al., 2003), the acceptance of gesture technology implied by the positive results (cf. **Figure 8-1**) can overestimate its true acceptance. On a more general level, all generalization made from the non-probabilistic sampling applied in this research to a broader population of older adults, even only to the matched age-band, must be treated with caution. Rather, the informative value of this research should be confined to an educated, motivated, and rather technology-minded subset of older adults. Further research is needed to gradually expand or delimitate the scope of these results to subgroups amongst the elderly, e.g. through stratified sampling.

8.2.2 Confounding factors

Prior Experience. Four of the five empirical studies have investigated group differences between an older and younger user group. Despite the fact that few differences between the age groups were found, the observed differences could have been caused by other factors such as the general affinity for technology or prior experience with similar systems. Previous research has already documented the influence of both an aging effect and an experience effect on the interaction performance with technical devices (Langdon, et al., 2007), and has highlighted the difficulty of disentangling these effects, since lower levels of technology experience are usually found at higher ages (Olson, et al., 2011). The user experience with a certain technology at hand is influenced by the technology experience throughout whole the lifetime (and especially during the formative period, Docampo Rama, 2001), as well as by age-related changes in sensory, cognitive and motor resources, and immediate prior interaction experience with related technology.

The present research cannot disentangle their relative impact on the outcome of the studies. However, we can argue that the observed effects are not predominantly

related to prior experience and refute the argument that age is negligible. Correlation analysis has shown that age group has not only the most and strongest correlations with gesture performance, but the correlations remain significant if various parameters of prior experience (mobile phones, touchscreens, gesture interaction) are accounted for (cf. Section 6.5.5). Conversely, if age is controlled for, the previously significant correlations of prior experience on gesture interaction disappear. Moreover, if prior experience parameters are entered as covariates in the analysis of variance of different input paradigms (gesture vs. keys), only partial evidence of an experience influence was observed (cf. Section 7.4.5). More importantly, prior touchscreen and gesture experience did not influence performance in gesture and key mode differentially. In contrast to Langdon and colleagues (2007), who found a dominating influence of prior experience, the present results hint at a moderate influence, and at the need to consider also other age-related factors.

Attitude towards technology. In this context, two further confounding factors can be discussed. Group differences between the younger and older adults might not be rooted in different levels of skills and knowledge, but explained rather by different attitudes towards technology, or different strategies during interaction. Attitude toward using technology was tested with the help of the questionnaire for technology affinity (Karrer, et al., 2009), including the subscales of *enthusiasm*, *competence*, *negative effects* and *positive effects*.

This measure was assessed as a control variable in studies 1A and 1B, but none of the subscale comparisons between the younger and older group was significant. We can therefore assume that performance differences in the gesture creation task were not primarily related to an increased enthusiasm, higher perceived competence, or differential expectations of positive and negative effects of technology usage in one of the age groups. These results are in line with the major body of research which could not substantiate a difference in attitudes towards computers between younger and older adults (Kelley & Charness, 1995). Furthermore, prominent models of technology acceptance also assume that attitude toward using technology does not have a significant influence on the intention to use technology (Venkatesh, et al., 2003).

Strategy: Speed-Accuracy Tradeoff. The gesture performance tasks (studies 1A & 1B) were carried out with different speed but comparable accuracy between the younger and older age group. The results have been discussed (see Section 5.6.2) in the light of a speed-accuracy tradeoff. It is well documented in literature that older users tend to sacrifice speed for an increase in accuracy (Salthouse, 1985; Welford, 1981). However, such a strategy must be understood as an adaptive process emerging during interaction, rather than a predisposition. Both age groups were controlled for their generalized action strategy (Nachtwei, 2008a), i.e. their general tendency to

prefer either speedy or accurate task fulfillment, but no group differences in action strategy could be substantiated.

8.2.3 Methodological Limitations

The classical problem of aging research is the confounding of age, cohort and period (Glenn, 2005). Traditional research designs all face the problem that two effects are always confounded: cross-sectional designs confound age and cohort, longitudinal designs confound age and period, while time-lag comparisons confound cohort and period (Rentz & Reynolds, 1981). Cohort analysis (any study in which one or more cohorts are measured at two or more points in time, Glenn, 2005) has been suggested as a group of methods to separate age, cohort and period effects, for example methods by Schaie (1965), Baltes (1968), Mason, Mason, Winsborough and Poole (1973) or Palmore (1978). However, most of the methods proposed for cohort analysis are based on additional assumptions (e.g. that effects are additive, or unidirectional and positive), and, more importantly, require elaborate research plans. Such a research plan, including multiple groups and multiple measurement points over an extended period of time, was not feasible with the given resources and time frame of our research.

From the above mentioned traditional research designs, time-lag designs did not apply since age-effects were of interest rather than generation (or cohort) effects. Longitudinal designs appeared problematic for various reasons. First, the limited time-frame of the project. Second, the known issues regarding selective sampling, selective survival, selective drop-out, or testing effects, which apply primarily to longitudinal designs (Baltes, 1968). Third, the consequences of repeated measures analysis were not desired, since the dependent variables were most interesting under the condition of first-time usage: how do older adults perform with gesture interaction even if they lack prior experience with such a system? Fourth, the rapid technological development could have confounded a longitudinal investigation. Over the time-course of this project, gesture technology has not only considerably matured (e.g. regarding the smoothness of interaction), it has also spread to a vast number of commercially available devices and is familiar nowadays to a much larger proportion of the population.

The only reasonable and feasible research approach appeared to be a cross-sectional comparison of different age groups at a specific point in time.

Cross-sectional design. In this design, age and generation effects are confounded, and – as discussed earlier – it is not legitimate to interpret obtained group differences as pure age effects (Baltes, 1968). This means that, apart from individual differences in prior experience with technology, the cohort of elderly users also share some

common experiences shaped by the global circumstances during their life, such as disruptive political or economic events, and incisive developments in science and technology. On these grounds Docampo Rama (2001) has formulated his concept of *technology generations* (see Section 2.1.2). In his taxonomy, our sample of older users would be classified as *early technical generation*. This cohort was confronted in their formative years mainly with electro-mechanical interfaces such as VCRs, televisions, radios etc., while the younger comparison group interacted in their formative years with computers, mobile phones and the internet. In the present research framework the degree to which these technology interaction experiences influences the interaction with a gesture-controlled device today (cohort effect) cannot be separated from the degree to which age-related physiological and psychological changes influence the interaction (age effect).

However, the applied focus of this research puts this problem into perspective. The focus of the work is centered around questions like *Can they use it? How do they use it?* and *How must we design so that they can use it?* rather than *What are the reasons behind their ability or inability to use it?* This work must thus be understood in the tradition of HCI research, rather than aging research. To this extent, experimental manipulations (besides grouping according to age) were of technical nature, instead of building subgroups according to cognitive performance or motor skills, for example. Taken together, the cross-sectional design with its inherent confounding of age and generation effects is not perceived as a major drawback given the research purpose. Nevertheless, caution is required with the causal interpretation of observed effects and generalization to other age-bands and cohorts.

Quasi-experimental setup. Since random assignment of subjects to experimental conditions is impossible, cross-sectional designs do not qualify for true experimental designs (Baltes, 1968; Schaie, 1977) and one has to resort to quasi-experimental designs (Baltes, 1968; Campbell & Stanley, 1963). The denomination of cross-sectional designs as quasi-experimental designs is not applied consistently in the field: while some literature defines quasi-experimental approaches on the basis of a non-random assignment of subjects to groups (explicitly including natural groups, as defined through age or gender, e.g. Bortz, 2005), others restrict quasi-experiments to situations where the cause is theoretically manipulable, excluding nonmanipulable causes such as age or sex (Shadish, Cook, & Campbell, 2002).

In our understanding the cross-sectional studies presented above are instances of quasi-experimental research designs. The well-known problems concerning internal and external validity (Campbell & Stanley, 1963) remain, and consequently alternative explanations cannot be ruled out as easily as with experimental designs (Shadish, et al., 2002). As discussed earlier, we cannot assume the participants' age as single cause of the observed effects: we can only state that after controlling for e.g.

computer literacy, mobile phone and touchscreen experience, or technology affinity, age remains a plausible explanation, especially if this is backed up by related research. Moreover, we put forward the earlier argument with respect to the cohort and age confounding: from an applied perspective, there is still value in observing age group differences in technology interaction, even without unambiguous determination of their cause. At the very least, the inferences drawn from the manipulated variables (e.g. screen size, gesture complexity, device posture or input method) are of value for the design of appropriate touch & gesture controlled devices.

Laboratory conditions. All studies have been conducted under laboratory conditions. Since the quasi-experimental design in combination with volunteer sampling already introduced an increasing number of potential alternative explanations, the laboratory setting was chosen to avoid a further decrease of internal validity. Moreover, studies 1A, 2A and 2B necessitated substantial computing equipment which would have rendered it impracticable to install in the field, e.g. in the older participants' homes. One can further argue that the artificial setups of pattern recreation (studies 1A and 1B) and gesture creation (studies 2A and 2B) principally lacked external validity (e.g. compared with the purposeful use of gesture to control a mobile device in context), so that carrying out the same study in the field would not have contributed to a substantial increase of external validity, but might have come at the cost of losing internal validity. The laboratory setting could at least minimize environmental effects, such as influences from noise, lighting conditions, posture, and reduce the risk of distractions.

It would be of great interest to see if the benefits of gesture control observed in the lab also manifest in the field. For example, different evaluation criteria could be applied, if participants have to interact unsupervised, while moving, or in a public and noisy environment. Future field research still has to demonstrate that the findings presented here can be generalized to real contexts and usage scenarios.

8.2.4 Limitations of the Gesture Set and Prototype

In Chapter 3 a classification system for gestures in HCI was proposed (cf. **Figure 3-1 G**) using 3 dimensions: movement complexity, concreteness of object relation, and gesturing space. While the tested gestures covered the dimension of *concreteness of object relation* (symbolic, iconic, indexical and manipulative gestures), *movement complexity* was restricted to one-finger gestures, and the *gesturing space* was restricted to a two-dimensional touch-sensitive surface. Thus, only a part of the possible gesture interaction space was covered by the present empirical investigations. Two questions follow: First: Were the tested gestures representative, so that results can be transferred to similar one-finger, surface gesture systems? And second: To what

extent, if at all, can the results be transferred to other gesture-based interfaces, which include for example multi-finger gestures, hand gestures on larger surfaces, or three dimensional gesture systems?

Validity of the gesture set. The gesture set used in studies 1A and 1B was compiled with reference to the gesture alphabets suggested by technology leaders in the field, e.g. GestureWorks (2011), Apple (2009), Microsoft (2009), or Opera (2011). Those gestures have been created for use with computers, tablets and mobile devices. In studies 1A and 1B, gestures of the same, or highly similar, shapes and complexity were tested on a stationary touchscreen and a mobile device. Thus, a generalization of the findings to tablet-size systems seems plausible (novel devices on the market blur the boundary between palm size and tablet size systems, and many of these devices work by the same operating systems and controls).

The prototype used in study 3 included, in contrast to most commercially available gesture controlled mobile devices, a relatively high proportion of iconic and symbolic gestures. However, the implemented patterns can also be found in commercially available systems. For example, the *arrowhead up* gesture is used in Sony's GPS navigation system, the *arrowhead left* is included in Microsoft's gesture recognizer (Microsoft, 2009), and the *question mark* is available in the Synaptics Gesture Suite (Synaptics Inc., 2010). *Alphanumeric symbols* have not only been employed in dedicated handwriting recognition systems like Graffiti (Palm Inc., 2012), but also as shortcut commands, e.g. to specify search functions (Purcher, 2012). The findings from the prototype evaluation can thus be assumed to be relevant for current and future mobile devices, both for systems relying only on multitouch and direct manipulation gestures, as well as for systems that incorporate symbolic input. To conclude, there is reason to believe that the results obtained in this research can be generalized to a broader range of mobile devices which use finger gesture input.

Multi-finger gestures were not tested on the functional prototype (study 3). The other studies, however, showed that gesture execution performance for two-finger gestures differs between younger and older adults with respect to error rate and accuracy (study 1B), and older adults have a general disinclination towards multi-finger gestures (studies 2A and 2B). The generally favorable results of gesture interaction for older users cannot therefore be transferred directly to multi-finger gestures. Further studies with functional devices are necessary to assess whether two- and more finger gestures are equally positively received by older users in real interaction settings. Furthermore, such studies need to control for prior experience with multitouch systems, which has been indicated to be a determinant of how intuitive these gestures are perceived (study 2A).

Generalization to other systems. A generalization of the results to larger surface gesturing systems (tables, walls), or even three-dimensional gesture control seems

not justified, since interaction is fundamentally different. For table-sized systems, substantially different gesture sets are proposed in the literature (Wobbrock, et al., 2009; Wu & Balakrishnan, 2003; Wu, Shen, Ryall, Forlines, & Balakrishnan, 2006), and sometimes large differences are observed between one-handed and bi-manual interaction (Frisch, Heyekorn, & Dachsel, 2009).

For three-dimensional gesturing in space, totally different gestures have to be devised, since the nature of interaction is fundamentally different compared with the drawing of surface gestures. Simple actions like selection, which can be realized effortlessly through a simple tap in surface gesturing, require new gestures e.g. the *AirTap* or *ThumbTrigger* (Fikkert, 2010). Importantly, in the context of older users, the psychomotor demands of three-dimensional gestures are high, since the lack of support of arm or wrist results in increased strain and fatigue.

Limitations to the generalization of the present results to similar devices due to the deficiencies of the functional prototype have been discussed at length in Section 7.4.3. The lack of multitouch capabilities, the non-optimal key layout, the restricted gesture detection, along with a screen design bias all contributed to put the positive evaluation of gesture input compared with key input into perspective. Future research needs to address these confounds, before a general verdict on the suitability of touch & gesture interaction for older adults can be returned.

8.3 Main Contributions

Gesture interaction had not previously been investigated systematically with regard to its usability for older adults. For this reason, the research described in this dissertation was largely explorative in nature. No theoretical accounts of the influence of age on gesture interaction performance exist, making it hard to interpret the results within a larger research framework. Similarly, the lack of empirical evidence in the field renders it difficult to anchor the observed results and put them into perspective. Rather, the presented results are seen as trailblazers of future studies in this domain, since the spreading of touch & gesture technology and demographic developments will require more attention to the central aim of this research in the future. The progress which has been made on a theoretical, methodological and empirical level is highlighted in the following sections.

8.3.1 Theoretical Contribution

The main theoretical contribution has been made in the understanding of gestural interfaces in human-computer interaction, irrespective of aging effects.

First, the characteristics of human gesture that bridge the various existing definitions and taxonomies in literature have been made explicit.

Second, based on a critical assessment of the different definitions of gesture in HCI, a new working definition of gesture in HCI has been proposed which attempts to reconcile the different aspects that previous definitions have pointed out: *“A gesture is a coordinated and intended movement of body parts to achieve communication. The information which it contains is specified by the configuration of body parts, the speed and direction of movement, and must be interpretable by its receiver”*. This definition does not constrain input by a particular input device (in contrast to Saffer, 2008), it adds the need of a receiver capable of interpretation to the central idea of Kurtenbach and Hulteen (1990), and at the same time does not exclude taps (which Kurtenbach and Hulteen, by implication, do). Hummels and Stappers’ (1998) definition is extended by stressing the intentional and coordinative nature of the movement.

Third, a new classification system of gestures in HCI has been proposed. Six previously suggested classification schemes were juxtaposed, disclosing their substantial differences. Instead of anchoring certain gesture types in a rigid hierarchical taxonomy (Hummels, et al., 1998; Pavlovic, et al., 1997; Quek, 1995), the proposed classification describes gestural interfaces along three continuous dimensions. Previous taxonomies neither agreed on the naming and granularity of classes, nor on the allocation of gestures to categories. The first dimension of the proposed system classifies gestures according to the relation of signs with associated meanings, inspired by the semiotic distinction introduced by Peirce (1982). The second dimension accounts for the technical constraints in human-machine systems by anchoring a gesture system with respect to the degrees of freedom of the sensor space, ranging from 2-dimensional surface gestures to remotely sensed 3D gestures in space. The third dimension accounts for the biophysical properties of a gesture. The continuum ranges from single finger movements, over multi-finger gestures, to one-handed and bi-manual gesturing. The benefit of this dimensional account is that it does not postulate selective and mutually exclusive categories, but rather understands the categories as examples along a spectrum that partly overlaps, leaving space for intermediate or mixed classes.

A fourth theoretical advance was made regarding the formalized description of gestures. In between unstructured natural language descriptions, and formalized description tools used in automated gesture recognition and synthesis, a semi-formal gesture description language was proposed (see Section 6.4.5). Based on a formalized

grammar which determines the basic structure of a well-formed gesture description and a (mostly) finite set of lexical entries, a verbal gesture description can be formed which is readable by non-experts. This middle ground of formalization and flexibility provides on the one hand enough uniformity for gestures to be described across researchers and projects with a standardized level of specification, but on the other hand leaves room to describe novel and creative gesture patterns due to the modular structure of the grammar. With the help of this gesture description language it was possible to file the movement patterns observed in study 2A into a catalogue containing over 500 user-derived gestures for 42 typical tasks on a mobile device. This gesture glossary could provide a useful source of inspiration for a system designer when considering alternative gesture implementations, or when insights about different conceptual models of users are asked for.

8.3.2 Methodological Contribution

From a methodological point of view, the major challenge was to establish this project at the middle ground of theory-driven and data-driven research agendas. A theoretical framework to contain gesture interaction specifically for elderly users was lacking. The closest theoretical frameworks that fitted the main research question were the various models of technology acceptance, some of them specified particularly for older users, as presented in Section 2.4 (Arning & Ziefle, 2007; Conci, et al., 2009; Renaud & van Biljon, 2008; van Biljon & Kotzé, 2007; Venkatesh & Davis, 2000; Venkatesh, et al., 2003). However, since their comparison revealed a lack of convergence, consistency, and a disregard of the actual abilities and limitations of elderly users, none of the models was selected to guide the present research. Instead, an applied perspective was taken that focuses on *actual use* rather than *intention to use*. To investigate actual use, a purely exploratory, data-driven approach could have been used, but this would have resulted only in post-hoc explanations with little chance to reveal small effects, and would have remained mostly on a descriptive level.

A mixture of theory-driven and data-driven investigations was therefore selected. The rich literature on motor control from aging psychology, combined with insights from HCI literature on the general benefits and drawbacks of input technologies for older adults, allowed to derive hypotheses for studies 1A and 1B, which could be translated into factorial designs and analyzed with the general linear model. Studies 2A and 2B, on the contrary, were purely explorative, since no reference points existed as to which type of gestures would be intuitive to older adults. Analysis was purely data driven, and the results remained mostly descriptive. Similarly, for study 3, no prior hypotheses could be derived concerning which input mode would be better received by the older participant group. Instead, based on the a posteriori observed

pattern of results, hypotheses can now be deduced (for example that older users are more efficient in accomplishing simple mobile phone tasks on gesture interfaces than on key interfaces) which have to be validated in future experiments.

Moreover, the plurality of measurement methods across the different studies contributed to a holistic account of the object of research. The studies in the first stage were primarily analyses on the basis of the behavioral data, which allowed assessing the movement trajectories objectively with temporal and geometric means. The measures of *form stability* and *directional stability* were of particular interest in this context, since they allowed for a common ground comparison of differently shaped gesture patterns. The usage of a paper-based mockup in study 2A allowed to open up the design space to a wealth of creative user input, but came at the cost of a laborious manual classification of video-recorded gestures by means of (subjective) visual inspection. The results of the video-based gesture evaluation in study 2B were predominantly obtained in form of subjective rating scales, rendering it easy to rank and quantify the results with respect to age group or gesture type. Finally, the design of a functional prototype allowed assessing the performance both through objective behavioral measures, such as interaction time and errors, and subjective data from usability questionnaires.

The strength of the presented approach lies in the combination of theory-driven and data-driven, of hypothesis-guided and explorative studies, in the plurality of different experimental setups (stationary as well as mobile), and the explanatory power obtained from the combination of objective and subjective measurements.

8.3.3 Empirical Contribution

The major contribution of this research must be seen in the advances toward an understanding how older adults handle touch and gesture-based interfaces. These results have been summarized in greater detail in Section 8.1. The main message emerging from the different empirical investigations is the insight that older adults can use gesture-based devices surprisingly well. The first research stage did not substantiate any particular difficulties of accurate gesture production among the older participants, even under conditions of small screen size, complex patterns or difficult posture. The second stage diagnosed highly similar preferences for suitable gestures for both old and young users, which indicates that no completely different gesture sets should be necessary for a younger and older user. On closer examination, certain preference patterns of older users span the studies: they are more willing to trade efficient execution of a pattern for familiarity, they are more tolerant towards iconic, symbolic and generally more complex patterns, and they show an increased disinclination with respect to curved patterns and multi-finger gestures. Most telling

are the results from the functional prototype evaluation. A direct comparison of gesture and key control in a mobile device revealed a consistent usability advantage of the gesture mode across a wide range of objective and subjective measurements.

With all befitted caution as a consequence of the limitations of the sample, the chosen research design, and the prototype features (cf. Section 8.2), the overall picture is encouraging. The existing literature on interface design for older adults has been extended by a first evaluation of the touch and gesture interaction which is prevailing on current mobile phones, and increasingly introduced in other devices. Overall, however, this new input paradigm does at least not seem to constitute a new barrier for older adults, or to keep them from adopting technology to a greater extent than previously available technology.

8.4 Practical Implications for Inclusive Gesture Control Design

General guidelines of how to design adequate products for older users have been suggested in the past, but mostly on a rather abstract level. For example, the guidelines published by the Center for Universal Design (e.g. Story, 2007) advise to accommodate flexible access and use, eliminate unnecessary complexity, provide consistency with user expectations, accommodate a wide range of literacy skills, reduce physical effort, minimize hazards, or provide fail safe features. While these principles apply to a wide range of interactive products and claim universal validity, they provide little practical guidance to the system designer. More specific design guidelines, including for example contrast ratios, font scaling, button spacings, color selection, auditory thresholds, moving distances etc. for a generally age-friendly HCI design can be found for example in Fisk and colleagues (2004), or Nichols and colleagues (2006). Murata and colleagues (2005) have already argued for the use of touch-panel interfaces particularly for older adults, and suggested a few design guidelines concerning target size, moving distance, locations and angles. However, since these recommendations were only established for simple taps on touch-only devices, design guidelines for *gesture control* systems which are adequate for older adults are missing so far.

The present research includes only a restricted number of empirical investigations and has limitations concerning the validity and the generalizability to other samples, devices or contexts (Section 8.2). Nevertheless, bearing these limitations in mind, the following guidelines for older users can be speculatively formulated, based on the empirical findings.

- Use touch & gesture input instead of key input for devices and tasks that are of complexity levels equal to standard mobile phone operations
- Consider gesture input for palm- and tablet sized devices; the gestures can in principal include multiple strokes, segments and fingers
- Accommodate for slower gesture execution for time-critical events (e.g. multi-tap latencies). Movement is slowed by 1.3 compared with younger users
- If possible, avoid multi-finger gestures
- Provide adequate training opportunities for users to familiarize them with the input technique and gesture patterns
- Employ curved gestures sparsely
- Prefer indexical and manipulative gestures over symbolic and iconic gestures
- Favor familiarity of a gesture pattern over its execution efficiency
- Include an easy way to return to the home screen
- Avoid alphanumerical gestures
- Provide strong visual aids to facilitate scrolling and avoid scrolling direction confusion
- If possible, include a question mark gesture for context-sensitive help

8.5 Future Directions

Saffer (2008, p. 3) claims that „tap is the new click“ and that with the advent of interactive gesture technology, „we’ve entered a new era of interaction design“. In a similar way, Widgor and Wixon (2011, p. 5) acknowledge that “we stand at the brink of another potential evolution in computing”. They argue that touch and gesture interaction – which they coin as “natural user interfaces” (NUIs) – “promise to reduce the barriers to computing still further, while simultaneously increasing the power of the user, and enabling computing to access still further niches of use”. Indeed, gesture technology has developed over the past years from a scattered use in top-of-the-line smartphones to a quasi-standard for most current mobile devices. Even beyond handheld systems, gesture control finds its way into desktop computing (e.g. the explicit tailoring of Windows 8 for touch control, Sinofsky, 2012), and even gesture control of cars is currently investigated (Burns, 2012). While time will tell whether users will really wave at their TVs and dashboards in the future, it is relatively safe to predict that gesture technology is here to stay, either to occupy certain product niches, or to dominate the whole computing landscape.

The present research has shown that this technological development does not bypass older users. On the contrary, the results indicate that the elderly do not exhibit serious difficulties during interaction with touch- and gesture control, but manage

interaction tasks even better under certain conditions. Moreover, they consider such input as intuitive to use, easily learnable and usable. In this light, it seems that touch and gesture control indeed bears the potential to bridge the “digital divide” that still exists between younger and older adults (Czaja & Hiltz, 2005). It goes without saying that this research presents only a first step towards a general conclusion about the suitability of gesture technology for older adults, and that more evidence is needed.

The next step should be a validation of the presented results using larger and more heterogeneous samples. Regarding the investigation of the relation of motor demand and gesture interaction performance (studies 1A and 1B), a replication should include standardized benchmark tests to assess motor performance (e.g. the Purdue Pegboard test, or the Motorische Leistungsserie (MLS), Schoppe, 1974). Furthermore, it would be interesting to investigate interaction parameters such as error rates and accuracy also within an even older sample, where hand movement restrictions become more common. To complement the present data, multi-finger gestures (which showed to be more difficult and less preferred by older users) should also be tested on larger, stationary screens, before a verdict on their general aptness is returned. Follow-up investigations on studies 2A & 2B should seek to separate the influences of prior experience from age-related effects more clearly. The idea of redundant gestures for the same action – e.g. a standard variant and a shortcut variant – could also be pursued further. Gesture evaluation on a functional prototype needs also be extended in order to validate the results of study 3. The identified technical issues (Section 7.4.3) need to be addressed, the sample enlarged, and the scope of tasks broadened. In addition, different hardware should be tested (e.g. a key interface which is better suited to older adults’ needs). Finally, the behavioral differences of a touch-interface compared with a key interface should be investigated separately to identify the differences which are due to gesturing as opposed to simple tapping.

On a more general account, the benefit of gestural interaction for older adults found in this research needs to be balanced against other input modalities, such as speech, or multimodal interaction. Previous research does not provide a univocal recommendation on which input modality is best for older adults (Naumann, et al., 2010; Weiss, Möller, & Schulz, 2012), since the modality choice is heavily modulated by task and the quality of interaction (e.g. the speech recognition rate). Nevertheless, as speech recognition becomes more common, more accurate, and slowly moves from command-based systems to natural speech interaction, voice input could position itself as even more senior-friendly input, since the verbal aspects of cognition are very well preserved over age. Therefore, future research in HCI for aging users should benchmark gesture interaction not only against traditional key input, but also against alternative emerging input styles, e.g. speech.

Taken together, a long and winding research path seems still ahead before a general conclusion on the aptness of gesture input for older adults can be reached. At this stage, the empirical findings need also to be fed back into theory and practice. For example, various models of technology usage and acceptance (Arning & Ziefle, 2007; Conci, et al., 2009; Renaud & van Biljon, 2008cf. Section 2.4) need to be revisited to accommodate the new insights generated by gesture interaction studies. In a similar way, design methodologies – such as *Inclusive Design*, or *Universal Design*, cf. Section 2.6 – need to be investigated in order to determine to which degree their principles support or discourage the use of touch and gesture interaction. Some guidelines (e.g. “use reasonable operating forces”, Story, 2007, p. 4) can be reconciled with gestural interfaces more easily than others (e.g. “facilitate the user’s accuracy and precision”). Lastly, these insights need also be anchored in the product development cycle. Already in the early phases of product development – planning, conceptualizing and primary layout – should the designer of a touch and gesture controlled device respect the design implications (Section 8.4) to accommodate the needs of older users. Only then will we see more senior-friendly touch and gesture interfaces on everyday technologies. Facilitating interaction will counter exclusion, counter technology anxiety, and help older adults to better master various activities in daily life.

References

- Aligna, J. (1982). Remarks on the analysis of covariance in repeated measures designs. *Multivariate Behavioral Research*, 17, 117-130.
- Allander, E., Björnsson, O. J., Ólafsson, Ó., Sigfússon, N., & Thorsteinsson, J. (1974). Normal range of joint movements in shoulder, hip, wrist and thumb with special reference to side: A comparison between two populations. *International Journal of Epidemiology*, 3(3), 253-261.
- Apple Inc. (2009). Multi-touch trackpad. Retrieved January 20, 2009, from <http://www.apple.com/macbook/features>
- Apple Inc. (2011a). Apple portables: Using the multi-touch glass trackpad. Retrieved September 4, 2011, from <http://support.apple.com/kb/ht3211>
- Apple Inc. (2011b). Multi-touch. Retrieved September 4, 2011, from <http://www.apple.com/iphone/design/>
- Apted, T., Kay, J., & Quigley, A. (2006). Tabletop sharing of digital photographs for the elderly. In *Proceedings of the ACM CHI 2006 conference on Human Factors in Computing Systems*. 781-790.
- Arias-Torres, D. (2006). The design and evaluation of a pen-based computer interface for novice older users. In P. Kellenberger (Ed.), *Proceedings of the 7th Mexican International Conference on Computer Science* (pp. 142-150). Los Alamitos, CA: IEEE Computer Society.
- Armbruster, C., Sutter, C., & Ziefle, M. (2007). Notebook input devices put to the age test: The usability of trackpoint and touchpad for middle-aged adults. *Ergonomics*, 50(3), 426-445.
- Armstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the nature of language*. Cambridge, UK: Cambridge University Press.
- Arning, K., Gaul, S., & Ziefle, M. (2010). "Same same but different" How service contexts of mobile technologies shape usage motives and barriers. In G. Leitner, M. Hitz & A. Holzinger (Eds.), *HCI in work and learning, life and leisure, LNCS Vol. 6389* (pp. 34-54). Berlin: Springer
- Arning, K., & Ziefle, M. (2007). Understanding age differences in PDA acceptance and performance. *Computers in Human Behavior*, 23(6), 2904-2927.
- Arning, K., & Ziefle, M. (2009). Different perspectives on technology acceptance: The role of technology type and age. In *Proceedings of the 5th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for e-Inclusion*. 20-41.

- Asano, Y., Saito, H., Sato, H., Wang, L., Gao, Q., & Rau, P.-L. P. (2007). Tips for designing mobile phone web pages for the elderly. In J. Jacko (Ed.), *Human-Computer Interaction, Part I, HCII 2007, LNCS Vol. 4450* (pp. 675-680). Berlin: Springer.
- Baddeley, A. (1992). Working Memory. *Science*, 255(5044), 556-559.
- Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. *Journal of the Association for Information Systems*, 8(4), 244-254.
- Baltes, P. B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development*, 11(3), 145-171.
- Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful Aging: Perspectives from the Behavioral Sciences* (pp. 1-34). New York: Cambridge University Press.
- Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology*, 50, 471-507.
- Becker, H. A. (2000). Discontinuous change and generational contracts. In S. Arber & C. Attias-Donfut (Eds.), *The myth of generational conflict: The family and state in ageing societies* (pp. 114-132). London: Routledge.
- Bergvall-Kåreborn, B., & Ståhlbrost, A. (2008). Participatory design: one step back or two steps forward? In *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*. 102-111.
- Bhuiyan, M., & Picking, R. (2009). Gesture-controlled user interfaces, what have we done and what's next? In U. Bleimann, P. S. Dowland, S. M. Furnell & V. M. Grout (Eds.), *Proceedings of the Fifth Collaborative Research Symposium on Security, E-learning, Internet and Networking (SEIN 2009)* (pp. 59-60). Plymouth, UK: University of Plymouth.
- Birren, J. E., & Schaie, K. W. (Eds.). (2001). *Handbook of the psychology of aging* (5th ed.). San Diego: Academic Press.
- Blackler, A., & Hurtienne, J. (2007). Towards a unified view of intuitive interaction: Definitions, models and tools across the world. *MMI Interaktiv*, 13, 36-54.
- Blackler, A., Popovic, V., & Mahar, D. (2002). Intuitive use of products. In D. Durling & J. Shackleton (Eds.), *Proceedings of the Design Research Society International Conference 2002* (pp. 120-134). London: Trentham Books.
- Blackler, A., Popovic, V., & Mahar, D. (2005). Intuitive interaction applied to interface design. In *Proceedings of the International Design Congress - IASDR '05, Douliou, Taiwan*.

- Bolt, R. A. (1980). "Put-that-there": Voice and gesture at the graphics interface. *Computer Graphics (ACM SIGGRAPH '80 Proceedings)*, 14(3), 262-270.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6th ed.). Heidelberg: Springer Medizin Verlag.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4th ed.). Heidelberg: Springer Medizin Verlag.
- Brooke, J. (1996). SUS - A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189-194). London: Taylor & Francis.
- Brown, M. (2008). The thumb generation. Retrieved November 28, 2011, from <http://www.boundlessline.org/2008/01/the-thumb-gener.html>
- Brown, R. (2007). fitellipse.m: Matlab Central, <http://www.mathworks.com/matlabcentral/fileexchange/15125-fitellipse-m>.
- Bruder, C. (2008). *Gestaltungsprinzipien für das Training älterer Benutzer elektronischer Geräte*. Unpublished doctoral dissertation, Technische Universität Berlin, Germany.
- Burns, C. (2012, Jan 27th). Mercedes-Benz DICE concept hands-on. Retrieved April 12, 2012, from <http://www.slashgear.com/mercedez-benz-dice-interior-concept-hands-on-27211028/>
- Buxton, W. (2011). Multi-touch systems that I have known and loved. Retrieved September 5, 2011, from <http://www.billbuxton.com/multitouchOverview.html>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1-22.
- Chaparro, A., Bohan, M., Fernandez, J., Choi, S. D., & Kattel, B. (1999). The impact of age on computer input device use: Psychophysical and physiological measures. *International Journal of Industrial Ergonomics*, 24(5), 503-513.
- Chaparro, A., Bohan, M., Fernandez, J., Kattel, B., & Choi, S. D. (1999). Is the trackball a better input device for the older computer user? *Journal of Occupational Rehabilitation*, 9(1), 33-43.
- Chaparro, A., Bohan, M., Scarlett, D. S., Fernandez, J., & Choi, S. D. (1999). The effect of age on computer input device use: Performance and ratings of perceived exertion. *SAE Technical Paper*, 1999-01-1604.

- Chaparro, A., Rogers, M., Fernandez, J., Bohan, M., Choi, S. D., & Stumpfhauser, L. (2000). Range of motion of the wrist: Implications for designing computer input devices for the elderly. *Disability & Rehabilitation*, 22(13), 633-637.
- Charness, N. (2008). Aging and human performance. *Human Factors*, 50(3), 548-555.
- Charness, N., & Boot, W. R. (2009). Aging and information technology use. Potentials and barriers. *Current Directions in Psychological Science*, 18(5), 253-258.
- Charness, N., Bosman, E. A., & Elliot, R. G. (1995). *Senior-friendly input devices: Is the pen mightier than the mouse?* Paper presented at the 103rd Annual Convention of the American Psychological Association, New York.
- Charness, N., Holley, P., Feddon, J., & Jastrzembski, T. (2004). Light pen use and practice minimize age and hand performance differences in pointing tasks. *Human Factors*, 46(3), 373-384.
- Charness, N., Schumann, C. E., & Boritz, G. M. (1992). Training older adults in word processing: Effect of age, training technique, and computer anxiety. *International Journal of Technology and Aging*, 5, 79-106.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conci, M., Pianesi, F., & Zancanaro, M. (2009). Useful, social and enjoyable: Mobile phone adoption by older people. In T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. A. Palanque, R. O. Prates & M. Winckler (Eds.), *Human-Computer Interaction - INTERACT 2009* (pp. 63-76). Berlin: Springer.
- Cooper, A. (2011, February 17). If users could lead innovation, they wouldn't be users. *Journal - A blog about design, business and the world we live in*. Retrieved February 12, 2012, from http://www.cooper.com/journal/2011/02/users_led_innovation.html
- Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., et al. (2006). Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (CREATE). *Psychology and Aging*, 21(2), 333-352.
- Czaja, S. J., & Hiltz, S. R. (2005). Digital aids for an aging society. *Communications of the ACM*, 48(10), 43-44.
- Czaja, S. J., & Lee, C. C. (2007). The impact of aging on access to technology. *Universal Access in the Information Society*, 5(4), 341-349.
- Czaja, S. J., & Moen, P. (2004). Technology and employment. In R. W. Pew & S. B. Van Hemmel (Eds.), *Technology for adaptive aging* (pp. 150-178). Washington, D.C.: The National Academies Press.
- Czaja, S. J., & Sharit, J. (1998). Age differences in attitudes toward computers. *Journal of Gerontology: Psychological Sciences*, 53B(5), 329-340.

- Davis, A. C., Ostri, B., & Parving, A. (1991). Longitudinal study of hearing. *Acta Otolaryngol*, 476, 12-22.
- Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Boston.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Delaney, H. D., & Maxwell, S. E. (1981). On using analysis of covariance in repeated measures designs. *Multivariate Behavioral Research*, 16, 105-123.
- Derwort, A. (1938). Untersuchungen über den Zeitverlauf figurierter Bewegungen beim Menschen. *Pflügers Archiv für die gesamte Physiologie*, 240, 661-675.
- Dienel, H.-L. (2007). Neues Leitbild. In W. Friesdorf & A. Heine (Eds.), *sentha - seniorengerechte Technik im häuslichen Alltag* (pp. 13-19). Berlin, Germany: Springer.
- Docampo Rama, M. (2001). *Technology generations handling complex user interfaces*. Unpublished doctoral dissertation, Eindhoven University of Technology, The Netherlands.
- Döring, T., Kern, D., Marshall, P., Pfeiffer, M., Schöning, J., Gruhn, V., et al. (2011). Gestural interaction on the steering wheel: Reducing the visual demand. In *Proceedings of the 2011 annual conference on Human Factors in Computing Systems*. 483-492.
- Dyck, J. L., Gee, N. R., & Al-Awar Smither, J. (1998). The changing construct of computer anxiety for younger and older adults. *Computers in Human Behavior*, 14(1), 61-77.
- Echtler, F., Klinker, G., & Butz, A. (2010). Towards a unified gesture description language. In *Proceedings of the 13th International Conference on Humans and Computers*. 177-182.
- Efron, D. (1941). *Gesture and environment*. New York: King's Crown Press.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1, 49-98.
- Ellis, R. D., & Allaire, J. C. (1999). Modeling computer interest in older adults: The role of age, education, computer knowledge, and computer anxiety. *Human Factors*, 41(3), 345-355.
- Federal Interagency Forum on Aging Related Statistics. (2010). *Older americans 2010: Key indicators of well-being*. Washington, DC.
- Fikkert, W. (2010). *Gesture interaction at a distance*. Enschede, Netherlands: Center for Telematics and Information Technology (CTIT).

- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley Publishing.
- Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2004). *Designing for older adults: Principles and creative human factors approaches*. Boca Raton, FL: CRC Press.
- fitage GmbH. (2006). Katharina das Große Handy zum Telefonieren. Retrieved August 28, 2011, from http://www.fitage.com/silver_age_technology_Katharina_das_Grosse/2,0,0,1,1.html
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391.
- Fozard, J. L. (1990). Vision and hearing in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 150-170). New York: Academic Press.
- Fozard, J. L., Vercruyssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R. E. (1994). Age differences and changes in reaction time: The Baltimore longitudinal study of aging. *Journal of Gerontology: Psychological Sciences*, 49(4), 179-189.
- Freeman, F. N. (1914). Experimental analysis of the writing movement. *Psychological Review Monographs Supplement*, 17, 1-46.
- Freudenthal, T. D. (1998). *Learning to use interactive devices: Age differences in the reasoning process*. Unpublished doctoral dissertation, Eindhoven University of Technology, The Netherlands.
- Friesen, W., Ekman, P., & Wallbott, H. (1979). Measuring hand movements. *Journal of Nonverbal Behavior*, 4(2), 97-112.
- Frisch, M., Heyekorn, J., & Dachsel, R. (2009). Investigating multi-touch and pen gestures for diagram editing on interactive surfaces. In G. Morrison & S. Subramanian (Eds.), *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (pp. 149-156). New York: ACM.
- Gardner, D. G., Discenza, R., & Dukes, R. L. (1993). The measurement of computer attitudes: An empirical comparison of available scales. *Journal of Educational Computing Research*, 9(4), 487-507.
- Gelderblom, H., Dyk, T. v., & van Biljon, J. (2010). Mobile phone adoption: Do existing models adequately capture the actual usage of older adults? In *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*. 67-74.
- GestureWorks. (2011). Open source gesture library. Retrieved October 20, 2011, from <http://gestureworks.com/features/open-source-gestures/>

- Giambra, L. M. (1997). Sustained attention and aging: Overcoming the decrement? *Experimental Aging Research*, 23(2), 145-161.
- Gibet, S., Kamp, J.-F., & Poirier, F. (2004). Gesture analysis: Invariant laws in movement. In A. Camurri & G. Volpe (Eds.), *Gesture-based communication in human-computer interaction* (pp. 451-452). Berlin, Germany: Springer.
- Glascocock, A. P., & Feinman, S. L. (1980). A holocultural analysis of old age. *Comparative Social Research*, 3, 311-332.
- Glenn, N. D. (2005). *Cohort analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Goebel, M. (2007). Ergonomic design of computerized devices for elderly persons - The challenge of matching antagonistic requirements. In C. Stephanidis (Ed.), *Universal access in human computer interaction. Coping with diversity* (pp. 894-903). Berlin, Germany: Springer.
- Greene, L. S., & Williams, H. G. (1996). Aging and coordination from the dynamic pattern perspective. In A. Ferrandez & N. Teasdale (Eds.), *Changes in sensory motor behavior in aging* (Vol. 114, pp. 89-131). New York, NY: Elsevier Science.
- Greenstein, J. S., & Arnaut, L. Y. (1988). Input devices. In M. Helander (Ed.), *The handbook of human-computer interaction* (pp. 495-516). Amsterdam: Elsevier.
- Gregor, P., & Newell, A. F. (2001). Designing for dynamic diversity: making accessible interfaces for older people. In *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly*. 90-92.
- Hanson, V. L. (2011). Technology skill and age: what will be the same 20 years from now? *Universal Access in the Information Society*, 10, 443-452.
- Hart, S. G. (2006). NASA-Taskload Index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*. 904-908.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 239-250). Amsterdam: North Holland Press.
- Hasegawa, S., Matsunuma, S., Omori, M., & Miyao, M. (2006). Aging effects on the visibility of graphic text on mobile phones. *Gerontechnology*, 4(4), 200-208.
- Hawthorn, D. (2000). Possible implications of aging for interface designers. *Interacting with Computers*, 12(5), 507-528.
- Herot, C., & Weinzapfel, G. (1978). One-point touch input of vector information from computer displays. *Computer Graphics*, 12(3), 210-216.
- Heuer, H. (1984). On re-scaleability of force and time in aiming movements. *Psychological Research*, 46(1), 73-86.

- Hirota, N. (2003). Reassessing current cell phone designs: Using thumb input effectively *CHI '03 extended abstracts on human factors in computing systems* (pp. 938-939). Ft. Lauderdale, FL, USA: ACM.
- Holzinger, A. (2002). User-centered interface design for disabled and elderly people: First experiences with designing a patient communication system (PACOSY). In K. Miesenberger, J. Klaus, W. Zagler & A. Karshmer (Eds.), *Computers helping people with special needs* (pp. 33-40). Berlin, Germany: Springer.
- Holzinger, A. (2003). Finger instead of mouse: Touch screens as a means of enhancing universal access. In N. Carbonell & C. Stephanidis (Eds.), *Universal access. Theoretical perspectives, practice, and experience* (pp. 387-397). Berlin, Germany: Springer.
- Homma, R. (2009). *Effects of the Wii on the physical and psychosocial condition of older adults in a senior residential facility*. The State University of New York, Buffalo, N.Y.
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107-129.
- Hummels, C., Smets, G., & Overbeeke, K. (1998). An intuitive two-handed gestural interface for computer supported product design. In *Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction*. 197-208.
- Hummels, C., & Stappers, P. J. (1998). Meaningful gestures for human computer interaction: Beyond hand postures. In *Proceedings of the IEEE International Conference on Automatic Face & Gesture Recognition FG'98*. 591-596.
- Hurtienne, J., & Blessing, L. (2007). Design for intuitive use - Testing image schema theory for user interface design. In J.-C. Bocquet (Ed.), *Proceedings of ICED '07 [CD-ROM]* (pp. 1-12). Paris: Ecole Centrale
- Hurtienne, J., & Naumann, A. (2010). QUESI - A questionnaire for measuring the subjective consequences of intuitive use. In R. Porzel, N. Sebanz & M. Spitzer (Eds.), *Interdisciplinary College 2010: Focus Theme: Play, Act and Learn* (pp. 536). Sankt Augustin: Fraunhofer Gesellschaft.
- Igbaria, M., & Chakrabarti, A. (1990). Computer anxiety and attitudes towards microcomputer use. *Behaviour & Information Technology*, 9(3), 229-241.
- Irie, T., Matsunaga, K., & Nagano, Y. (2005). Universal design activities for mobile phone: Raku Raku PHONE. *Fujitsu Scientific and Technical Journal. Special Issue on Universal Design*, 41(1), 78-85.
- Jagacinski, R. J., Liao, M.-J., & Fayyad, E. A. (1995). Generalized slowing in sinusoidal tracking by older adults. *Psychology and Aging*, 10(1), 8-19.
- Jané-Llopis, E., & Gabilondo, A. (Eds.). (2008). *Mental health in older people. Consensus Paper*. Luxembourg: European Communities.

- Jastrzemski, T., Charness, N., Holley, P., & Feddon, J. (2005). Input devices for web browsing: Age and hand effects. *Universal Access in the Information Society*, 4(1), 39-45.
- Jay, G. M., & Willis, S. L. (1992). Influence of direct computer experience on older adults' attitudes toward computers. *Journal of Gerontology: Psychological Sciences*, 47(4), 250-257.
- Jerger, J. (1973). Audiological findings in aging. *Advances in Otorhinolaryngology*, 20, 115-124.
- Jin, Z., Plocher, T., & Kiff, L. (2007). Touch screen user interfaces for older adults: Button size and spacing. In C. Stephanidis (Ed.), *Universal access in human computer interaction. Coping with diversity* (pp. 933-941). Berlin, Germany: Springer.
- Johnson, E. A. (1965). Touch Display - A novel input/output device for computers. *Electronics Letters*, 1(8), 219-220.
- Johnson, E. A. (1967). Touch displays: A programmed man-machine interface. *Ergonomics*, 10(2), 271-277.
- Kane, S. K., Wobbrock, J. O., & Ladner, R. E. (2011). Usable gestures for blind people: understanding preference and performance *Proceedings of the 2011 annual conference on human factors in computing systems* (pp. 413-422). Vancouver, BC, Canada: ACM.
- Kanstrup, A. M., & Christiansen, E. (2006). Selecting and evoking innovators: Combining democracy and creativity. In *Proceedings of the 4th Nordic conference on Human-Computer Interaction: changing roles*. 321-330.
- Karam, M., Lee, J. C., Rose, T., Quek, F., & McCrickard, S. (2009). Comparing gesture and touch for notification system interactions. In *Proceedings of the Second International Conference on Advances in Computer-Human Interactions, ACHI '09*. 7-12.
- Karam, M., & Schraefel, M. C. (2005). *A taxonomy of gestures in human computer interaction*: Technical Report ECSTR-IAM05-009, Electronics and Computer Science, University of Southampton.
- Karlson, A. K., Bederson, B. B., & Contreras-Vidal, J. L. (2006a). *Studies in one-handed mobile design: Habit, desire and agility*. Tech report HCIL-2006-02: Computer Science Dept., University of Maryland, College Park, MD.
- Karlson, A. K., Bederson, B. B., & Contreras-Vidal, J. L. (2006b). Understanding one handed mobile device interaction. In J. Lumsden (Ed.), *Handbook of research on user interface design and evaluation for mobile technology* (pp. 86-101). London: IGI Global.

- Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009). Technikaffinität erfassen - Der Fragebogen TA-EG. In A. Lichtenstein, C. Stöbel & C. Clemens (Eds.), *Der Mensch als Mittelpunkt technischer Systeme. 8. Berliner Werkstatt Mensch-Maschine-Systeme* (pp. 196-201). Düsseldorf, Germany: VDI Verlag GmbH.
- Keates, S., & Clarkson, P. J. (2003). Countering design exclusion through inclusive design. In *Proceedings of the CUU '03, Vancouver, BC, Canada*.
- Kelley, C. L., & Charness, N. (1995). Issues in training older adults to use computers. *Behaviour & Information Technology*, 14(2), 107-120.
- Kendon, A. (1986). Current Issues in the study of gesture. In J.-L. Nespoulous, P. Perron & A. R. Lecours (Eds.), *Biological Foundations of Gesture* (pp. 23-47). Hillsdale, New Jersey: Lawrence Erlbaum, Associates.
- Kendon, A. (2000). Language and gesture: unity or duality? In D. McNeill (Ed.), *Language and Gesture* (pp. 47-63). Cambridge, UK: Cambridge University Press.
- Ketcham, C. J., Seidler, R. D., Van Gemmert, A. W., & Stelmach, G. E. (2002). Age-related kinematic differences as influenced by task difficulty, target size, and movement amplitude. *Journal of Gerontology: Psychological Sciences and Social Sciences*, 57(1), 54-64.
- Ketcham, C. J., & Stelmach, G. E. (2004). Movement control in the older adult. In P. R. W. & S. B. van Hemel (Eds.), *Technology for adaptive aging* (pp. 64-92). Washington, DC: The National Academies Press.
- Kim, H., Heo, J., Shim, J., Kim, M., Park, S., & Park, S. (2007). Contextual research on elderly users' needs for developing universal design mobile phone. In C. Stephanidis (Ed.), *Universal access in human computer interaction. Coping with diversity* (pp. 950-959). Berlin, Germany: Springer.
- Kirakowski, J. (1996). The Software Usability Measurement Inventory: background and usage. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 169-178). Brighton: Taylor & Francis.
- Kirk, H. (1992). Geriatric medicine and the categorisation of old age - the historical linkage. *Ageing and Society*, 12, 483-497.
- Kölsch, M., & Martell, C. (2006). Towards a common human gesture description language. In *Proceedings of the IEEE VR 2006 Workshop Specification of Mixed Reality User Interfaces*. 25-28.
- Krampe, R. T. (2002). Aging, expertise and fine motor movement. *Neuroscience & Biobehavioral Reviews*, 26(7), 769-776.
- Kranstedt, A., Kopp, S., & Wachsmuth, I. (2002). MURML: A multimodal utterance representation markup language for conversational agents. In *Proceedings of the Workshop Embodied Conversational Agents, Autonomous Agents and Multi-Agent Systems (AAMAS02)*.

- Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: a process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-281). Cambridge, UK: Cambridge University Press.
- Kray, C., Nesbitt, D., Dawson, J., & Rohs, M. (2010). User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proceedings of the 12th international conference on Human Computer Interaction with Mobile Devices and Services*. 239-248.
- Krueger, M. W., Gionfriddo, T., & Hinrichsen, K. (1985). VIDEOPLACE -- An artificial reality. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI'85)*. 35-40.
- Kühnel, C., Westermann, T., Hemmert, F., Kratz, S., Müller, A., & Möller, S. (2011). I'm home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies*, 69(11), 693-704.
- Kurniawan, S. (2006). An exploratory study of how older women use mobile phones. In *Proceedings of the 8th International conference on Ubiquitous Computing*. 105-122.
- Kurniawan, S. (2008). Older people and mobile phones: A multi-method investigation. *International Journal of Human-Computer Studies*, 66(12), 889-901.
- Kurniawan, S., Mahmud, M., & Nugroho, Y. (2006). A study of the use of mobile phones by older persons. In G. Olson (Ed.), *CHI '06 extended abstracts on Human Factors in Computing Systems* (pp. 989-994). New York: ACM.
- Kurtenbach, G., & Hulteen, E. A. (1990). Gestures in human-computer communication. In B. Laurel (Ed.), *The art of human computer interface design* (pp. 309-317). Boston, MA: Addison-Wesley.
- Labovitz, S. (1970). The assignment of numbers to rank order categories. *The American Sociological Review*, 35(3), 515-524.
- Laguna, K., & Babcock, R. L. (1997). Computer anxiety in young and older adults: Implications for human-computer interactions in older populations. *Computers in Human Behavior*, 13(3), 317-326.
- Langdon, P., Lewis, T., & Clarkson, J. (2007). The effects of prior experience on the use of consumer products. *Universal Access in the Information Society*, 6(2), 179-191.
- LaVictoire, M., & Everhart, N. (2009). A touch screen button size and spacing study with older adults. In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction. Addressing Diversity, LNCS Vol. 5614* (pp. 261-262). Berlin: Springer.
- Lee, S.-S., Kim, S., Jin, B., Choi, E., Kim, B., Jia, X., et al. (2010). How users manipulate deformable displays as input devices. In *Proceedings of the 28th international conference on Human Factors in Computing Systems (CHI'10)*. 1647-1656.

- Lee, Y. S. (2007). *Older adult's user experiences with mobile phones: identification of user clusters and user requirements*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Leonardi, C., Albertini, A., Pianesi, F., & Zancanaro, M. (2010). An exploratory study of a touch-based gestural interface for elderly. In E. Hvannberg & M. Lárusdóttir (Eds.), *Proceedings of NordiCHI '10* (pp. 845-850). New York: ACM.
- Leonardi, C., Mennecozzi, C., Not, E., Pianesi, F., & Zancanaro, M. (2008). Designing a familiar technology for elderly people. *Gerontechnology*, 7(2), 151.
- Liao, M.-J., Jagacinski, R. J., & Greenberg, N. (1997). Quantifying the performance limitations of older and younger adults in a target acquisition task. *Journal of Experimental Psychology: Human Perception and Performance*, 23(6), 1644-1664.
- Lin, C., Hsieh, T.-L., & Shiang, W.-J. (2009). Exploring the interface design of mobile phone for the elderly. In M. Kurosu (Ed.), *Human centered design, LNCS Vol. 5619* (pp. 476-481). Berlin: Springer.
- Loewenfeld, I. E. (1979). Pupillary changes related to age. In H. S. Thompson, R. Daroff, L. Frisen, J. S. Glaser & M. D. Sanders (Eds.), *Topics in Neuro-Ophthalmology* (pp. 124-150). Baltimore: Williams & Wilkins.
- Losson, O., & Vannobel, J.-M. (1999). Sign specification and synthesis. In *Proceedings of the International Gesture Workshop on Gesture-Based Communication in Human-Computer Interaction*. 239-251.
- Mace, R. L., Hardie, G. J., & Place, J. P. (1991). Accessible environments: Toward Universal Design. In W. E. Preiser, J. C. Vischer & E. T. White (Eds.), *Design intervention: Toward a more humane architecture*. New York: Van Nostrand Reinhold.
- Mahmud, M., & Kurniawan, H. (2005). Involving psychometric tests for input device evaluation with older people. In *Proceedings of the 17th Australian conference on Computer-Human Interaction (OZCHI '05)*. 1-10.
- Maier, T. (2008). Vom senioren gerechten Design zum 'Design for all' - Studien zum Produktdesign wider die Stigmatisierung. In *Proceedings of the Produktdesign für alle: Für Junge = Für Alte?*, 13-21.
- Marketing Charts. (2010). Young consumers want cell phone features. Retrieved October 15, 2011, from <http://www.marketingcharts.com/direct/young-consumers-want-cell-phone-features-13278/>
- Martin, M., & Kliegel, M. (2005). *Psychologische Grundlagen der Gerontologie*. Stuttgart, Germany: Kohlhammer.
- Mason, K. O., Mason, W. M., Winsborough, H. H., & Poole, W. K. (1973). Some methodological issues in cohort analysis of archival data. *American Sociological Review*, 38, 242-258.

- Mathiowetz, V., Kashman, N., Volland, G., Weber, K., Dowe, M., & Rogers, S. (1985). Grip and pinch strength: Normative data for adults. *Archives of Physical Medicine and Rehabilitation*, 66(2), 69-74.
- MathWorks. (2011). MATLAB - The language of technical computing. Retrieved December 1, 2011, from <http://www.mathworks.com/products/matlab/index.html>
- Mauney, D., Howarth, J., Wirtanen, A., & Capra, M. (2010). Cultural similarities and differences in user-defined gestures for touchscreen user interfaces. In *CHI '10 extended abstracts on Human Factors in Computing Systems* (pp. 4015-4020). Atlanta, Georgia, USA: ACM.
- Mauney, D., & Le Hong, S. (2010). Cultural differences and similarities in the use of gestures on touchscreen user interfaces. In *Proceedings of the Usability Professionals' Association International Conference (UPA '10)*. 1-16.
- Mawston, N. (2011). Apple becomes world's largest handset vendor by revenue in Q1 2011. Retrieved October 4, 2011, from <http://www.strategyanalytics.com/default.aspx?mod=reportabstractviewer&a0=6263>
- McDowd, J. M., & Craik, F. I. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 267-280.
- McDowd, J. M., & Shaw, R. J. (2000). Attention and aging: A functional perspective. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 221-292). Mahwah, NJ: Erlbaum.
- McLaughlin, A. C., Rogers, W. A., & Fisk, A. D. (2003). Effects of attentional demand on input device use in younger and older adults. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 47, 247-251.
- McLaughlin, A. C., Rogers, W. A., & Fisk, A. D. (2009). Using direct and indirect input devices: Attention demands and age-related differences. *ACM Transactions on Computer-Human Interaction*, 16(1), 1-15.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University Press Chicago.
- McNeill, D. (2000). Introduction. In D. McNeill (Ed.), *Language and gesture* (pp. 1-10). Cambridge, UK: Cambridge University Press.
- Mehta, N. (1982). *A flexible machine interface*. Unpublished Master Thesis, University of Toronto, Toronto.
- Meier, S. T. (1985). Computer aversion. *Computers in Human Behavior*, 1(2), 171-179.

- Melenhorst, A.-S., Rogers, W. A., & Bouwhuis, D. G. (2006). Older adults motivated choice for technological innovation: Evidence for benefit-driven selectivity. *Psychology and Aging*, 21(1), 190-195.
- Melenhorst, A.-S., Rogers, W. A., & Caylor, E. C. (2001). The use of communication technologies by older adults: Exploring the benefits from the user's perspective. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 45, 221-225.
- Microsoft. (2009). Application gestures and semantic behavior. Retrieved January 20, 2009, from <http://msdn.microsoft.com/en-us/library/windows/desktop/ms704830%28v=vs.85%29.aspx>
- Microsoft. (2011). Gestures: flick, pan, and stretch. Retrieved September 4, 2011, from <http://www.microsoft.com/windowsphone/en-us/howto/wp7/start/gestures-flick-pan-and-stretch.aspx>
- Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15(1), 31-49.
- Mitsubishi Electric Research Laboratories. (2011). DiamondTouch. Retrieved October 1, 2011, from <http://www.merl.com/areas/DiamondTouch/>
- Mitzner, T. L., Mayhorn, C. B., Fisk, A. D., & Rogers, W. A. (2006). Ergonomics and aging. In W. S. Marras & W. Karwowski (Eds.), *Fundamentals and assessment tools for occupational ergonomics* (pp. 1-22). Boca Raton, FL: CRC Press.
- Mohs, C., Hurtienne, J., Israel, J. H., Naumann, A., Kindsmüller, M. C., Meyer, H. A., et al. (2006). IUUI - Intuitive use of user interfaces. In T. Bosnick, M. Hassenzahl, M. Müller-Prove & M. Peissner (Eds.), *Usability professionals 2006* (pp. 130-133). Stuttgart, Germany: German Chapter der Usability Professionals Association e.V.
- Morris, C. W. (1938/1970). *Foundations of the theory of signs* (Vol. 6-7). Chicago: Chicago University Press.
- Murata, A., & Iwase, H. (2005). Usability of touch-panel interfaces for older adults. *Human Factors*, 47(4), 767-776.
- Murata, A., & Takahashi, R. (2008). Optimal slope of touch panel - Comparison between young and older adults. In *Proceedings of the Fourth International Workshop on Computational Intelligence & Applications: IW CIA 2008*. 111-116.
- Nachtwei, J. (2008a). *Ein experimenteller Ansatz in der Kognitiven Ergonomie: Personenmerkmale als Kodeterminanten der Komplexität einer belebten Mikrowelt*. Saarbrücken: Vdm Verlag.
- Nachtwei, J. (2008b). *Testhandbuch zum Fragebogen zur Erfassung von Strategie und Expertise in Experimenten, revidierte Fassung*. Berlin: Humboldt-Universität.

- Naumann, A., Hurtienne, J., Israel, J. H., Mohs, C., Kindsmüller, M. C., Meyer, H. A., et al. (2007). Intuitive use of user interfaces: Defining a vague concept. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (pp. 128-136). Berlin, Germany: Springer.
- Naumann, A., Wechsung, I., & Hurtienne, J. (2010). Multimodal interaction: A suitable strategy for including older users? *Interacting with Computers*, 22(6), 465-474.
- Neisser, U. (1963). Decision-time without reaction-time: experiments in visual scanning. *American Journal of Psychology*, 76(3), 376-385.
- Nichols, T. A., Rogers, W. A., & Fisk, A. D. (2006). Design for aging. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (pp. 1418-1445): John Wiley and Sons, Inc.
- Nintendo. (2011). About the Wii console. Retrieved October 16, 2011, from <http://www.nintendo.com/wii/console>
- Nintendo Co. Ltd. (2011). Consolidated Sales Transition by Region. Retrieved August 29, 2011, from http://www.nintendo.co.jp/ir/library/historical_data/pdf/consolidated_sales_e1106.pdf
- Nitz, J. C., Kuys, S., Isles, R., & Fu, S. (2010). Is the Wii Fit™ a new-generation tool for improving balance, health and well being? A pilot study. *Climateric*, 13(5), 487-491.
- Norman, D. A., & Nielsen, J. (2010). Gestural interfaces: A step backwards in usability. *Interactions*, 17(3), 46-49.
- Nysveen, H., Pedersen, P., & Thorbjørnsen, H. (2005). Intentions to use mobile services: Antecedents and cross-service comparisons. *Journal of the Academy of Marketing Science*, 33(3), 330-346.
- O'Hanlon, A., Bond, R., Knapp, B., & Carragher, L. (2010). The nestling project: Attitudes toward technology and associations with health, relationships, and quality of life. *Gerontechnology*, 9(2), 236.
- Oehl, M., Sutter, C., & Ziefle, M. (2007). Considerations on efficient touch interfaces - How display size influences the performance in an applied pointing task. In M. J. Smith & G. Salvendy (Eds.), *Human interface and the management of information. Methods, techniques and tools in information design* (pp. 136-143). Berlin, Germany: Springer.
- Olson, K. E., O'Brien, M. A., Rogers, W. A., & Charness, N. (2011). Diffusion of technology: Frequency of use for younger and older adults. *Ageing International*, 36(1), 123-145.

- Olwal, A., Lachanas, D., & Zacharouli, E. (2011). OldGen: Mobile phone personalization for older adults. In *Proceedings of the CHI 2011 (SIGCHI Conference on Human Factors in Computing Systems)*. 3393-3396.
- Opera Software ASA. (2011). Mouse gestures in Opera. Retrieved October 20, 2011, from <http://www.opera.com/browser/tutorials/gestures/>
- Osman, Z., Maguire, M., & Tarkiainen, M. (2003). Older users' requirements for location based services and mobile phones. In L. Chittaro (Ed.), *Human-computer interaction with mobile devices and services* (pp. 352-357). Berlin, Germany: Springer.
- Otjacques, B., Krier, M., Feltz, F., Ferring, D., & Hoffmann, M. (2009). *User-centered activity management system for elderly people*. Paper presented at the 3rd International Conference on Pervasive Computing Technologies for Healthcare 2009 (Pervasive Health '09), London, UK.
- Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research*, 23(7), 689-699.
- Pak, R., McLaughlin, A. C., Lin, C.-C., Rogers, W. A., & Fisk, A. D. (2002). An age-related comparison of a touchscreen and a novel input device. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 46(2), 189-192.
- Palm Inc. (2012). Das Graffiti 2-Schreibsystem verwenden. Retrieved February 12, 2012, from http://kb.hpwebos.com/wps/portal/kb/common/article/8366_de.html#1
- Palmore, E. (1978). When can age, period and cohort be separated? *Social Forces*, 57, 282-295.
- Pattison, M., & Stedmon, A. (2006). Inclusive design and human factors: Designing mobile phones for older users. *PsychNology Journal*, 4(3), 267-284.
- Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997). Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 677-695.
- Peirce, C. S. (1982). On a new list of categories (1867). In Peirce Edition Project (Ed.), *Writings of Charles S. Peirce: A chronological edition* (Vol. W2. 56). Indianapolis: Indianapolis University Press.
- Persad, U., Langdon, P., & Clarkson, P. J. (2006). Inclusive design evaluation and the capability-demand relationship. In P. J. Clarkson, P. M. Langdon & P. Robinson (Eds.), *Designing accessible technology* (pp. 177-188). London: Springer.
- Petras, A. (2006). Wertorientierte Vermarktungsstrategien in der Best Ager-Zielgruppe. *Planung & Analyse*, Vol. 2/2006.

- Pew Internet & American Life Project. (2010). Data sets may 2010 – Cell phones, Pew Research Center's Internet & American Life Project. Retrieved June 22, 2011, from <http://www.pewinternet.org/Shared-Content/Data-Sets/2010/May-2010--Cell-Phones.aspx>
- Phang, C. W., Sutanto, J., Kankanhalli, A., Li, Y., Tan, B. C. Y., & Teo, H.-H. (2006). Senior citizens' acceptance of information systems: A study in the context of e-government services. *IEEE Transactions on Engineering Management*, 53(4), 555-569.
- PirkI, J. J. (1994). *Transgenerational Design: Products for an aging population*. New York: John Wiley & Sons.
- Pohlmeyer, A., Blessing, L., Wandke, H., & Maue, J. (2009). The value of answers without question[s]: A qualitative approach to user experience and aging. In M. Kurosu (Ed.), *Human Centered Design, HCII 2009, LNCS Vol. 5619* (pp. 894-903).
- Pohlmeyer, A., Machens, F., & Blessing, L. (2010). Attractive or not - what's the difference? Inter- and intra-group comparisons in the Kano model. *Proceedings of the 11th International Design Conference - DESIGN 2010*, 413-422.
- Psychtoolbox.org. (2011). Psychtoolbox-3 Retrieved December 1, 2011, from <http://psychtoolbox.org/HomePage>
- Purcher, J. (2012, February 24). Google invents an original search gesture for future devices. *Patently Apple* Retrieved February 28, 2012, from <http://www.patentlyapple.com/patently-apple/2012/02/google-invents-an-original-search-gesture-for-future-devices.html>
- Quek, F. (1995). Eyes in the interface. *International Journal of Image and Vision Computing*, 13(6), 511-525.
- Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., et al. (2002). Multimodal human discourse: gesture and speech. *ACM Transactions on Computer-Human Interaction*, 9(3), 171-193.
- Rabbitt, P. (1965). An age-decrement in the ability to ignore irrelevant information. *Journal of Gerontology*, 20, 233-238.
- Regan, D. T., & Fazio, R. (1977). On the consistency between attitudes and behavior: Look to the method of attitude formation. *Journal of Experimental Social Psychology*, 13(1), 28-45.
- Renaud, K., & van Biljon, J. (2008). Predicting technology acceptance and adoption by the elderly: A qualitative study. In *Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: Riding the wave of technology*. 210-219.

- Rentz, J. O., & Reynolds, F. D. (1981). Separating age, cohort and period effects in consumer behavior. *Advances in Consumer Research*, 8, 596-601.
- Rhyne, J. (1987). Dialogue management for gestural interfaces. *Computer Graphics*, 21(2), 137-142.
- Riviere, C. N., & Thakor, N. V. (1996). Effects of age and disability on tracking tasks with a computer mouse: Accuracy and linearity. *Journal of Rehabilitation Research and Development*, 33(1), 6-15.
- Rogers, W. A., Fisk, A. D., McLaughlin, A. C., & Pak, R. (2005). Touch a screen or turn a knob: Choosing the best device for the job. *Human Factors*, 47(2), 271-288.
- Rogers, W. A., O'Brien, M. A., & McLaughlin, A. C. (2006). Selection and design of input devices for assistive technologies. In *Proceedings of the 9th International Conference on Control, Automation, Robotics, and Vision (ICARCV '06)*. 92-97.
- Rohrmann, B. (1978). Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. *Zeitschrift für Sozialpsychologie*, 9, 222-245.
- Ruiz, J., Li, Y., & Lank, E. (2011). User-defined motion gestures for mobile interaction. In *Proceedings of the SIGCHI 2011 annual conference on Human Factors in Computing Systems (CHI'11)*. 197-206.
- Sackmann, R., & Weymann, A. (1994). *Die Technisierung des Alltags - Generationen und technische Innovation*. Frankfurt/Main: Campus Verlag.
- Saffer, D. (2008). *Designing gestural interfaces*. Sebastopol, CA: O'Reilly Media.
- Salhouse, T. A. (1985). Speed of behavior and its implications for cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging (2nd Ed.)* (pp. 400-426). New York: Van Nostrand.
- Salhouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103(3), 403-428.
- Salhouse, T. A., & Ferrer-Caja, E. (2003). What needs to be explained to account for age-related effects on multiple cognitive variables? *Psychology and Aging*, 18(1), 91-110.
- Salhouse, T. A., Mitchell, D. R. D., Skovronek, E., & Babcock, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15(3), 507-516.
- Sarlegna, F. R. (2006). Impairment of online control of reaching movements with aging: A double-step study. *Neuroscience Letters*, 403, 309-314.
- Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, 64, 92-107.

- Schaie, K. W. (1977). Quasi-experimental research designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 39-58). New York: Van Nostrand Reinhold.
- Schaie, K. W. (1996). Intellectual development in adulthood. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 266-286). San Diego: Academic Press.
- Schaie, K. W. (2004). Cognitive aging. In R. W. Pew & S. B. van Hemel (Eds.), *Technology for adaptive aging* (pp. 43-63). Washington, DC: The National Academies Press.
- Schieber, F. (2003). Human factors and aging: Identifying and compensating for age-related deficits in sensory and cognitive function. In N. Charness & K. W. Schaie (Eds.), *Impact of technology on successful aging* (pp. 42-84). New York: Springer.
- Schieber, F., & Kline, D. W. (1994). Age differences in the legibility of symbol highway signs as a function of luminance and glare level: A preliminary report. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 38(2), 133-136.
- Schneider, N., Schreiber, S., Wilkes, J., Grandt, M., & Schlick, C. M. (2008). Foundations of an age-differentiated adaptation of the human-computer interface. *Behaviour & Information Technology*, 27(4), 319-324.
- Schnock, A. (2011). *Visuelles Feedback für eine intuitive Gestensteuerung mobiler Endgeräte*. Unpublished master thesis, Technische Universität Berlin, Berlin.
- Schoppe, K. J. (1974). Das MLS-Gerät: ein neuer Testapparat zur Messung feinmotorischer Leistungen. *Diagnostica*, 20, 43-47.
- Schröder, S., & Ziefle, M. (2006). Icon design on small screens: Effects of miniaturization on speed and accuracy in visual search. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 656-660.
- Schröder, S., & Ziefle, M. (2008). Effects of icon concreteness and complexity on semantic transparency: Younger vs. older users. In K. Miesenberger, J. Klaus, W. Zagler & A. Karshmer (Eds.), *Proceedings of the 11th International Conference on Computers Helping People with Special Needs (ICCHP '08)*, LNCS Vol. 5105 (pp. 90-97). Berlin: Springer.
- Schroeder, D. H., & Salthouse, T. A. (2004). Age-related effects on cognition between 20 and 50 years of age. *Personality and Individual Differences*, 36(2), 393-404.
- Schuler, D., & Namioka, A. (1993). Preface. In D. Schuler & A. Namioka (Eds.), *Participatory Design: Principles and practices* (pp. xi-xiii). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Scialfa, C. T. (2002). The role of sensory factors in cognitive aging research. *Canadian Journal of Experimental Psychology*, 56(3), 153-163.

- Sengpiel, M., & Dittberner, D. (2008). The computer literacy scale (CLS) for older adults - Development and validation. In M. Herczeg & M. C. Kindsmüller (Eds.), *Mensch & Computer 2008: Viel Mehr Interaktion* (pp. 7-16). Lübeck, Germany: Oldenbourg.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, New York: Houghton Mifflin.
- Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. *Computer*, 16(8), 57-69.
- Siek, K. A., Rogers, Y., & Connelly, K. H. (2005). Fat finger worries: How older and younger users physically interact with PDAs. In M. F. Costabile & F. Paternò (Eds.), *Human-Computer Interaction - INTERACT 2005* (pp. 267-280). New York: Springer.
- Silverstone, R., & Haddon, L. (1996). Design and the domestication of ICTs. Technical change and everyday life. In R. Silverstone & R. Mansell (Eds.), *Communication by design. The politics of information and communication technologies*. Oxford: Oxford University Press.
- Sinofsky, S. (2012, March 28th). Touch hardware and Windows 8. Retrieved April 12, 2012, from <http://blogs.msdn.com/b/b8/archive/2012/03/28/touch-hardware-and-windows-8.aspx>
- Smith, A. (2011). 35% of American adults own a smartphone. Pew Internet and American Life Project. Retrieved August 28, 2011, from <http://pewinternet.org/Reports/2011/Smartphones.aspx>
- Smith, M. W., Sharit, J., & Czaja, S. J. (1999). Aging, motor control, and the performance of computer mouse tasks. *Human Factors*, 41(3), 389-396.
- Somberg, B. L. (1986). A comparison of rule-based and positionally constant arrangements of computer menu items. *SIGCHI Bull.*, 18(4), 255-260.
- Spirduso, W. W., & Choi, J. (1993). Age and practice effects on force control of the thumb and index fingers in precision pinch and bilateral coordination. In G. E. Stelmach & V. Hûmberg (Eds.), *Sensorimotor impairments in the elderly: Are they reversible?* (pp. 393-412). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Statistisches Bundesamt. (2011). *Internationale Bildungsindikatoren im Ländervergleich*. Wiesbaden, Germany: Statistisches Bundesamt.
- Statistisches Bundesamt Deutschland. (2006). Bevölkerung Deutschlands bis 2050 - Ergebnisse der 11. koordinierten Bevölkerungsvorausberechnung Variante 1-W1. Retrieved August 18, 2010, from [https://www-ec.destatis.de/csp/shop/sfg/bpm.html.cms.cBroker.cls?cmspath=struktur,vollanzeige.csp&ID=1019440](https://www.ec.destatis.de/csp/shop/sfg/bpm.html.cms.cBroker.cls?cmspath=struktur,vollanzeige.csp&ID=1019440)

- Stone, R. G. (2008). Mobile touch interfaces for the elderly. In G. Bradley (Ed.), *Proceedings of ICT, Society and Human Beings 2008*. Amsterdam: International Association for Development of the Information Society (IADIS).
- Story, M. F. (2007). Applying the principles of Universal Design to medical devices. In J. M. Winters & M. F. Story (Eds.), *Medical instrumentation - Accessibility and usability considerations* (pp. 83-93). Boca Raton, FL: CRC Press.
- Stößel, C., & Blessing, L. (2009). *Is gesture-based interaction a way to make interfaces more intuitive and accessible?* Paper presented at the 23rd BCS conference on Human Computer Interaction (HCI '09), Cambridge, United Kingdom.
- Stößel, C., & Blessing, L. (2010a). Mobile device interaction gestures for older users. In E. ó. Hvannberg, M. K. Lárusdóttir, A. Blandford & J. Gulliksen (Eds.), *Proceedings of the 6th Nordic Conference on Human-Computer Interaction* (pp. 793-796). Reykjavik, Iceland ACM.
- Stößel, C., & Blessing, L. (2010b). Tap, swipe & pinch: Designing suitable multi-touch gestures for older users *Proceedings of the 11th International DESIGN Conference* (pp. 463-472). Dubrovnik, Croatia: Design Society.
- Stößel, C., Wandke, H., & Blessing, L. (2009). An evaluation of finger-gesture interaction on mobile devices for elderly users. In A. Liechtenstein, C. Stößel & C. Clemens (Eds.), *Prospektive Gestaltung von Mensch-Technik-Interaktion. 8. Berliner Werkstatt Mensch-Maschine-Systeme* (pp. 470-475). Düsseldorf, Germany: VDI.
- Stößel, C., Wandke, H., & Blessing, L. (2010). Gestural interfaces for elderly users: Help or hindrance? In S. Kopp & I. Wachsmuth (Eds.), *Gesture in Embodied Communication and Human-Computer Interaction, LNCS Vol. 5934* (pp. 269-280). Berlin: Springer.
- Sülzenbrück, S., Hegele, M., Heuer, H., & Rinkenauer, G. (2010). Generalized slowing is not that general in older adults: Evidence from a tracing task. *Occupational Ergonomics*, 9(2), 111-117.
- Sutter, C., & Müsseler, J. (2007). User specific design of interfaces and interaction techniques: What do older computer users need? In C. Stephanidis (Ed.), *Universal access in human computer interaction. Coping with diversity* (pp. 1020-1029). Berlin, Germany: Springer.
- Swinnen, S. P. (1998). Age-related deficits in motor learning and differences in feedback processing during the production of a bimanual coordination pattern. *Cognitive Neuropsychology*, 15(5), 439-466.
- Synaptics Inc. (2010). Synaptics previews Scribe- The revolutionary next generation in PC touch interface. Retrieved February 12, 2012, from <http://www.synaptics.com/about/press/press-releases/synaptics-previews-scribe-revolutionary-next-generation-pc-touch-interfac>

- Tang, H.-H., & Kao, S.-A. (2005a). Understanding the real need of the elderly people when using mobile phones. *Proceedings of include '05 [CD-ROM]*. London: Royal College of Art.
- Tang, H.-H., & Kao, S.-A. (2005b). Understanding the user's model of the elderly people while using mobile phones. In H. R. Arabnia (Ed.), *Proceedings of the 2005 HCII '05 [CD-ROM]*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tang, H.-H., & Peng, S.-Y. (2006). *Holistically understanding the user model of the elderly people while using mobile phones*. Paper presented at the 7th Asia-Pacific Conference on Computer-Human Interaction (APCHI'06), Taipei, Taiwan.
- Taveira, A. D., & Choi, S. D. (2009). Review study of computer input devices and older users. *International Journal of Human-Computer Interaction*, 25(5), 455-474.
- Thalmeier, S., & Koller, F. (2009). Multi-Touch-Produkte: Welche kulturellen Unterschiede und Gemeinsamkeiten gibt es bei der Nutzung von Gesten? *Usability Professionals Special Issue*, 79-83.
- Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*, 52(2), 336-358.
- Tränkle, U., & Deutschmann, D. (1991). Factors influencing speed and precision of cursor positioning using a mouse. *Ergonomics*, 34(2), 161-174.
- Tremblay, F., Wong, K., Sanderson, R., & Côté, L. (2003). Tactile spatial acuity in elderly persons: Assessment with grating domes and relationship with manual dexterity. *Somatosensory & Motor Research*, 20(2), 127-132.
- Tull, D. S., & Hawkins, D. I. (1993). *Marketing research: Measurement and method*. New York: Macmillan Publishing.
- United Nations. (2009). Population ageing and development 2009. Retrieved August 2, 2011, from <http://www.un.org/esa/population/publications/ageing/ageing2009chart.pdf>
- United Nations. (2010). World population prospects, the 2010 revision. Retrieved August 2, 2011, from http://esa.un.org/unpd/wpp/Sorting-Tables/tab-sorting_ageing.htm
- van Biljon, J., & Kotzé, P. (2007). Modelling the factors that influence mobile phone adoption. In *Proceedings of the 2007 Annual research conference of the South African institute of computer scientists and information technologists (SAICIST 2007)*. 152-161.
- van Biljon, J., & Kotzé, P. (2008). Cultural factors in a mobile phone adoption and usage model. *Journal of Universal Computer Science*, 14(16), 2650-2679.

- Vandercammen, M. (2007). De senioren - Typologie. Retrieved August 2, 2011, from <http://www.oivo-crioc.org/files/nl/2088nl.pdf>
- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186-204.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.
- Vienna Institute of Demography. (2010). European demographic data sheet 2010. Retrieved August 2, 2011, from http://www.oeaw.ac.at/vid/datasheet/download/European_Demographic_Data_Sheet_2010.pdf
- Walker, J. (1990). Through the looking glass. In B. Laurel (Ed.), *The art of human computer interface design* (pp. 439-447). Reading, MA: Addison Wesley.
- Walker, N., Philbin, D. A., & Fisk, A. D. (1997). Age-related differences in movement control: Adjusting submovement structure to optimize performance. *Journal of Gerontology: Psychological Sciences*, 52(1), 40-52.
- Weale, R. A. (1961). Retinal illumination and age. *Transactions of the Illuminating Engineering Society*, 26, 95-100.
- Weiss, B., Möller, S., & Schulz, M. (2012). Modality preferences of different user groups. In *Proceedings of the The Fifth Interational Conference on Advances in Computer-Human Interations (ACHI 2012)*. 354-359.
- Welford, A. T. (1958). *Aging and human skill. A report centred on work by the Nuffield Unit for research into problems of aging*. Oxford: Oxford University Press.
- Welford, A. T. (1977). Motor Performance. In J. E. Birren & K. W. Schaie (Eds.), *Handbook for the psychology of aging* (pp. 450-496). New York: Van Nostrand Reinhold.
- Welford, A. T. (1981). Signal, noise, performance and age. *Human Factors*, 23(1), 97-109.
- Welford, A. T., Norris, A. H., & Shock, N. W. (1969). Speed and accuracy of movement and their changes with age. *Acta Psychologica*, 30(0), 3-15.
- Wexelblat, A. D. (1994). *A feature-based approach to continuous-gesture analysis*. Unpublished master thesis, Massachusetts Insitute of Technology, Boston.
- WHO (World Health Organization). (2010). Definition of an older or elderly person. Retrieved 2. August, 2011, from <http://www.who.int/healthinfo/survey/ageingdefnolder/en/index.html>
- Wickremaratchi, M. M., & Llewelyn, J. G. (2006). Effects of ageing on touch. *Postgraduate Medical Journal*, 82(967), 301-304.

- Widgor, D., & Wixon, D. (2011). *Brave NUI World: designing natural user interfaces for touch and gesture*. Burlington, MA, USA: Morgan Kaufmann.
- Wilkowska, W., & Ziefle, M. (2009). Which factors form older adults' acceptance of mobile information and communication technologies? In A. Holzinger & K. Miesenberger (Eds.), *HCI and Usability for e-Inclusion, LNCS Vol. 5889* (pp. 81-101). Berlin: Springer.
- Wobbrock, J. O., Aung, H. H., Rothrock, B., & Myers, B. A. (2005). Maximizing the guessability of symbolic input. In G. C. van der Veer & C. Gale (Eds.), *CHI '05 extended abstracts on Human Factors in Computing Systems* (pp. 1869-1872). New York: ACM.
- Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. In S. Greenberg, S. E. Hudson, K. Hinkley, M. Ringel Morris & D. R. Olsen (Eds.), *CHI '09: Proceedings of the 27th International Conference on Human Factors in Computing Systems* (pp. 1083-1092). New York: ACM.
- Wolf, K., Dicke, C., & Grasset, R. (2011). Touching the void: Gestures for auditory interfaces. In *Proceedings of the Fifth international conference on tangible, embedded, and embodied interaction*. 305-308.
- Wood, E., Willoughby, T., Rushing, A., Bechtel, L., & Gilbert, J. (2005). Use of computer input devices by older adults. *Journal of Applied Gerontology*, 24(5), 419-438.
- Worcester, R. M., & Burns, T. R. (1975). A statistical examination of the relative precision of verbal scales. *Journal of Market Research Society*, 17(2), 181-197.
- Wromblewski, L. (2010). Touch gesture reference guide. Retrieved January 14, 2012, from <http://www.lukew.com/ff/entry.asp?1071>
- Wu, M., & Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User Interface Software and Technology*. 193-202.
- Wu, M., Shen, C., Ryall, K., Forlines, C., & Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, Washington, DC*. 185-192.
- Yoshioka, K. (2005). *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse*. Unpublished doctoral dissertation, Rijksuniversiteit Groningen, Groningen, NL.
- Zickuhr, K. (2010, December 16). Generations 2010. Pew Internet & American Life Project. Retrieved August 28, 2011, from <http://www.pewinternet.org/Reports/2010/Generations-2010.aspx>

- Ziefle, M., & Bay, S. (2004). Mental models of a cellular phone menu. Comparing older and younger novice users. In S. Brewster & M. Dunlop (Eds.), *Mobile Human-Computer Interaction - MobileHCI '04* (pp. 25-37). Berlin, Germany: Springer.
- Ziefle, M., & Bay, S. (2005). How older adults meet complexity: Aging effects on the usability of different mobile phones. *Behaviour and Information Technology*, 24(5), 375-389.
- Ziefle, M., & Bay, S. (2006). How to overcome disorientation in mobile phone menus: A comparison of two different types of navigation aids. *Human-Computer Interaction*, 21(4), 393-433.

List of Abbreviations

ANOVA	Analysis Of Variance (statistical procedure)
ATM	Automated teller machine (cash machine)
BI	Behavioral Intention to Use
CA	Computer Anxiety
CE	Computer Efficacy
<i>df</i>	degrees of freedom (statistical parameter)
ERE	Extended Regular Expression
GUI	Graphical User Interface
HCI	Human-Computer Interaction
ID	Inclusive Design
K-S	Kolmogorov-Smirnov test (statistical procedure)
<i>M</i>	mean (statistical parameter)
<i>Mdn</i>	median (statistical parameter)
MOPTAM	Mobile Phone Technology Acceptance Model
<i>N</i>	sample size (statistical parameter)
NUI	Natural User Interface
<i>p</i>	probability (statistical parameter)
PD	Participatory Design
PDA	Personal Digital Assistant
PEOU	Perceived Ease of Use
PU	Perceived Usefulness
PS	Perceived Safety
QUESI	QUEstionnaire on the Subjective consequences of Intuitive use
<i>SD</i>	standard deviation (statistical parameter)
SOC	Selection, Optimization & Compensation
STAM	Senior Technology Acceptance Model
STC	Subjective technical confidence
SUMI	Software Usability Measurement Inventory (questionnaire)
SUS	System Usability Scale (questionnaire)
TA-EG	TechnikAffinität bei Elektronischen Geräten (questionnaire)
TAM	Technology Acceptance Model
TAM2	revised version of TAM
TFT	Thin-film transistor (technology used in liquid crystal displays)
TRA	Theory of Reasoned Action
<i>U</i>	Mann-Whitney U (statistical parameter)
UCD	User Centered Design
UD	Universal Design
UI	User Interface
UTAUT	Unified Theory of Acceptance and Use of Technology
χ^2	chi-square (statistical parameter)
<i>z</i>	z-score (statistical parameter)

Appendix

A.1 Overview of Mobile Interaction Problems

Table A.1-1. Overview of common problems of older adults during mobile phone interaction

Category	Problem	Sources
Buttons	Buttons are too small	(Kim, et al., 2007; Kurniawan, 2008; Kurniawan, et al., 2006; Tang & Kao, 2005b; Tang & Peng, 2006)
	Buttons are too closely arranged	
	Buttons provide no “click” feedback when pressed	
	Short press and long press cannot be discerned	
	Buttons with too much label information (e.g. number and several letters associated to the same button)	
	Unlabeled buttons: unclear functionality	
	Soft keys: button representing multiple functions depending on context causes confusion	
Text reading	Font size too small ; particularly on small screen devices, and even with corrective lenses	(Asano, et al., 2007; Hasegawa, Matsunuma, Omori, & Miyao, 2006; Kim, et al., 2007; Kurniawan, 2008; Kurniawan, et al., 2006; Tang & Kao, 2005b; Tang & Peng, 2006)
	Difficult if contrast between text and background is low	
	Difficult if font is not gothic	
	Abbreviated and foreign words are skipped	
Text entry	General difficulties with text entry	(Asano, et al., 2007; Kim, et al., 2007; Kurniawan, 2008; Kurniawan, et al., 2006)
	Predictive texting is distracting and confusing: deleting wrong predictions demands multiple key presses to delete	
	Timing difficulties: cursor moves on if the interval between two presses is too long: character is missed	
Functions	Too many functions cause confusion	(Kim, et al., 2007; Kurniawan, 2008; Kurniawan, et al., 2006)
	Many functions irrelevant and unnecessary	
	Too many steps necessary for goal achievement	
Symbols & Icons	Icon comprehension difficult in general Icon preferences sometimes differ between young	(Asano, et al., 2007; Kim, et al., 2007; Schröder & Ziefle,

	and old users	2008; Tang & Kao, 2005b;
	Especially small status icons (e.g. time, battery life) often misunderstood	Tang & Peng, 2006)
	Scrolling and blinking objects are skipped (not enough affordance)	
Focus recognition	Focus areas too small	(Asano, et al., 2007)
	Focus highlighting based on color (e.g. hyperlinks) not easily discernible	
Menu navigation	Generally low navigation performance, especially at high complexities	(Kim, et al., 2007; Tang & Kao, 2005a; Tang & Peng, 2006; Ziefle & Bay, 2004)
	Mental models often reflect a shallow menu structure	
	Difficulties to remember and retrieve functions in deep menus	
	Mental models of older users often inadequate to grasp hierarchical nature of menus	
Conceptual	Automatic wizards pose difficulties	(Asano, et al., 2007; Kim, et al., 2007; Tang & Kao, 2005b; Tang & Peng, 2006)
	Meaning of missed calls sometimes unclear	
	Difficulties understanding differences between similar functions, e.g. calling from phone book and calling from history	
	Misunderstanding of color red as always indicating warning	
	Mode uncertainties for similar screens, e.g. shooting and album mode within the camera function	
Device	Size sometimes too small to hold comfortably	(Kim, et al., 2007; Kurniawan, 2008; Kurniawan, et al., 2006)
	Ring tone not loud enough (and lack of knowledge how to change)	
	Speaker volume often too low (and lack of knowledge of volume control)	
	Camera lens gets blocked by fingers	
	Camera out of focus because device is shaken too much during photography	

A.2 Study 1A: Gesture Performance on a Stationary Screen

A.2.1 Questionnaire Data

Table A.2-1. Comparison of **age groups** on the NASA-TLX, TA-EG and FESE-R instruments.
Results of the Mann-Whitney Test

instrument	subscale	older		younger		<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
NASA-TLX	Mental Demand	5.82	4.22	5.00	2.58	149.00	-0.133	.902	-.02
	Physical Demand*	4.50	3.36	7.47	4.27	93.00	-1.997	.046	-.34
TA-EG	Enthusiasm	3.06	1.14	2.90	1.13	142.00	-0.634	.536	-.11
	Competence	3.38	0.51	3.19	0.41	121.00	-1.315	.194	-.22
	Positive effects	3.78	0.55	3.60	0.59	128.00	-1.084	.284	-.18
	Negative effects	2.60	0.74	7.74	0.46	139.00	-0.734	.473	-.12
FESE-R	Action strategy	4.59	0.76	4.24	0.61	100.00	-1.962	.051	-.33

A.2.2 Gesture Performance Analysis

Table A.2-2. Results of the Three-way GLM Repeated Measures Analysis on **ERROR RATE**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η</i> ²
Age group	1,34	1.253	.271	.036
Size ***	1.43, 48.53	12.940	.000	.276
Complexity	2,68	1.140	.325	.032
Age group x Size	1.43, 48.53	3.143	.068	.085
Age group x Complexity	2,68	0.685	.507	.020
Size x Complexity	4,136	0.943	.441	.027
Age group x Size x Complexity	4,136	0.737	.568	.021

Table A.2-3. Results of the Three-way GLM Repeated Measures Analysis on **BOUNDARY VIOLATIONS**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part. η^2</i>
Age group	1,34	0.020	.888	.001
Size ***	1.27,43.05	12.964	.000	.276
Complexity	2,68	0.487	.617	.014
Age group x Size	1.27,43.05	0.136	.873	.004
Age group x Complexity	2,68	1.482	.234	.042
Size x Complexity *	3.07,104.34	2.747	.045	.075
Age group x Size x Complexity	3.07,104.34	0.402	.756	.012

Table A.2-4. Results of the Three-way GLM Repeated Measures Analysis on **AVERAGE VELOCITY**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part. η^2</i>
Age group **	1,34	6.671	.014	.164
Size ***	1.17, 37.99	340.325	.000	.909
Complexity ***	1.38,46.83	52.375	.000	.606
Age group x Size ***	1.17, 37.99	12.139	.001	.263
Age group x Complexity	1.38,46.83	0.454	.565	.013
Size x Complexity ***	2.79,94.70	25.238	.000	.426
Age group x Size x Complexity	2.79,94.70	2.195	.098	.061

Table A.2-5. Results of the Three-way GLM Repeated Measures Analysis on **ANGULAR DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part. η^2</i>
Age group	1,34	0.006	.938	.000
Size ***	2,68	232.502	.000	.872
Complexity ***	1.66,56.58	83.652	.000	.711
Age group x Size	2,68	0.709	.496	.020
Age group x Complexity	1.66,56.58	0.584	.531	.017
Size x Complexity	2.96,100.77	2.480	.066	.068
Age group x Size x Complexity	2.96,100.77	0.716	.543	.021

Table A.2-6. Results of the Two-way GLM Repeated Measures Analysis on **ECCENTRICITY**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,34	1.030	.317	.029
Size ***	2,68	60.694	.000	.641
Age group x Size	2,68	0.465	.630	.013

Table A.2-7. Results of the Three-way GLM Repeated Measures Analysis on **LINE DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,34	0,399	.598	.009
Size ***	1.37,46.42	232.635	.000	.872
Complexity	2,68	2.942	.063	.080
Age group x Size **	1.37,46.42	5.536	.015	.140
Age group x Complexity	2,68	1.402	.253	.040
Size x Complexity ***	4,136	4.064	.004	.107
Age group x Size x Complexity ***	4,136	5.459	.000	.138

Table A.2-8. Results of the Two-way GLM Repeated Measures Analysis on **ELLIPSE DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,34	0.229	.635	.007
Size ***	2,68	155.440	.000	.821
Age group x Size	2,68	1.416	.250	.040

A.3 Study 1B: Gesture Performance on a Mobile Device

A.3.1 Sample Characteristics

Table A.3-1. Comparison of **age groups** on the NASA-TLX, TA-EG and FESE-R instruments.
Results of the Mann-Whitney Test. MD =mental demand, PD = physical demand

instrument	subscale	older		younger		U	z	p	r
		M	SD	M	SD				
NASA-TLX	MD (one finger)	7.08	3.78	5.34	2.30	131.00	-1.455	.149	-.24
	PD (one finger)	7.05	3.96	7.03	3.72	176.50	-0.117	.914	-.02
	MD (two finger)	7.95	4.07	7.32	3.55	161.50	-0.559	.586	-.09
	PD (two finger)	8.63	3.46	8.68	3.76	171.50	-0.264	.800	-.04
TA-EG	Enthusiasm	2.94	0.80	2.99	0.99	178.50	-0.059	.960	-.01
	Competence	3.08	0.57	3.34	0.45	131.00	-1.471	.145	-.24
	Positive effects	3.56	0.53	3.46	0.72	173.50	-0.206	.844	-.03
	Negative effects	2.97	0.88	3.02	0.89	176.50	-0.117	.914	-.02
FESE-R	Action strategy	4.66	0.71	4.36	0.53	128.50	-1.519	.132	-.25

A.3.2 Gesture Performance Analysis: One Finger Gestures

Table A.3-2. Comparison of dependent variables according to different levels of *posture* (table, hand thumb) and *difficulty* (easy, medium, hard).

size:		table		hand		thumb	
		M	SD	M	SD	M	SD
error rate	(%)	1.45	3.15	0.83	1.11	2.69	2.50
average velocity	(cm/s)	9.10	2.88	8.95	2.86	6.02	2.10
angular deviation	(°deg)	5.53	1.47	5.31	1.40	7.61	1.67
eccentricity		0.57	0.07	0.57	0.08	0.60	0.06
line deviation	(mm)	0.32	0.13	0.29	0.10	0.59	0.13

complexity:		easy		medium		hard	
		M	SD	M	SD	M	SD
error rate	(%)	-	-	-	-	-	-
average velocity	(cm/s)	8.22	2.68	7.66	2.36	8.19	2.29
angular deviation	(°deg)	5.64	1.24	6.55	2.25	6.26	1.30
eccentricity		-	-	0.64	0.07	0.53	0.05
line deviation	(mm)	0.44	0.11	0.35	0.09	0.41	0.09

Table A.3-3. Results of the Three-way GLM Repeated Measures Analysis on **VELOCITY**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group ***	1,35	9.022	.005	.205
Difficulty ***	1.60,56.02	26.097	.000	.427
Posture ***	2,70	97.949	.000	.737
Age group x Difficulty	1.60,56.02	0.725	.460	.020
Age group x Posture	2,70	0.068	.912	.002
Difficulty x Posture ***	2.52,88.06	9.618	.000	.216
Age group x Difficulty x Posture	2.52,88.06	0.642	.563	.018

Table A.3-4. Results of the Three-way GLM Repeated Measures Analysis on **ANGULAR DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,35	1.068	.308	.030
Difficulty *	2,70	4.258	.018	.108
Posture ***	2,70	52.476	.000	.600
Age group x Difficulty	2,70	0.420	.658	.012
Age group x Posture	2,70	0.088	.915	.003
Difficulty x Posture	4,140	2.100	.084	.057
Age group x Difficulty x Posture	4,140	1.515	.201	.041

Table A.3-5. Results of the Three-way GLM Repeated Measures Analysis on **LINE DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,35	0.225	.638	.006
Difficulty ***	2,70	70.307	.000	.668
Posture ***	1.48,51.88	117.632	.000	.771
Age group x Difficulty	2,70	2.206	.118	.059
Age group x Posture	1.48,51.88	0.173	.776	.005
Difficulty x Posture	2.76,96.59	2.101	.110	.057
Age group x Difficulty x Posture	2.76,96.59	1.631	.170	.045

Table A.3-6. Results of the Three-way GLM Repeated Measures Analysis on **ECCENTRICITY**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,35	0.348	.559	.010
Difficulty ***	1,35	132.659	.000	.791
Posture **	2,70	4.409	.016	.112
Age group x Difficulty	1,35	2.547	.119	.068
Age group x Posture	2,70	0.031	.970	.001
Difficulty x Posture	2,70	2.070	.134	.056
Age group x Difficulty x Posture	2,70	2.268	.111	.061

Table A.3-7. Results of the Three-way GLM Repeated Measures Analysis on **ELLIPSE DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group *	1,35	4.786	.035	.120
Difficulty ***	1,35	259.306	.000	.881
Posture *	1.48,51.71	4.804	.020	.121
Age group x Difficulty	1,35	0.021	.881	.001
Age group x Posture	1.48,51.71	1.152	.310	.032
Difficulty x Posture	1.68,58.94	0.451	.606	.013
Age group x Difficulty x Posture	1.68,58.94	1.150	.316	.032

Table A.3-8. Results of the Two-way GLM Repeated Measures Analysis on **ERROR RATE**.
(Error rate was not differentiated between difficulty levels)

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group *	1,34	6.712	.014	.165
Posture **	1.58,53.59	7.509	.003	.181
Age group x Posture	1.58,53.59	1.142	.316	.032

Table A.3-9. **Age group** differences of gesture performance parameters.
Results of the Mann-Whitney test. (p-values for exact two-tailed comparison)

Dependent variable	older		younger		<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Error rate ***	2.37	1.92	1.01	1.19	79.50	-2.618	.008	-.44
Angular deviation	6.38	1.48	5.95	1.01	132.00	-1.185	.245	-.19
Line deviation	0.41	0.09	0.39	0.10	152.00	-0.577	.578	-.09
Ellipse deviation	0.59	0.13	0.51	0.09	107.00	-1.945	.053	-.32

Table A.3-10. Friedman's ANOVA comparing effects of **posture** for older and younger adults.
For dependent variables which are non-normal distributed in one or both of the age groups

Dependent variable	older			younger		
	<i>df</i>	χ^2	<i>p</i>	<i>df</i>	χ^2	<i>p</i>
Error rate	2	18.178	.000	2	16.545	.000
Angular deviation	2	19.444	.000	2	28.526	.000
Line deviation	2	24.778	.000	2	29.158	.000
Ellipse deviation	2	14.111	.001	2	12.316	.001

Table A.3-11. Friedman's ANOVA comparing effects of **difficulty** for older and younger adults
For dependent variables which are non-normal distributed in one or both of the age groups
(Error rate was not differentiated between difficulty levels)

Dependent variable	older			younger		
	<i>df</i>	χ^2	<i>p</i>	<i>df</i>	χ^2	<i>p</i>
Angular deviation	2	4.111	.144	2	6.000	.053
Line deviation	2	21.000	.000	2	32.316	.000
Ellipse deviation	1	14.222	.000	1	19.000	.000

A.3.3 Gesture Performance Analysis: Two Finger Gestures

Table A.3-12. Results of the Two-way GLM Repeated Measures Analysis on **ERROR RATE**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group *	1,36	5.670	.023	.136
Posture ***	1,36	14.260	.001	.284
Age group x Posture	1,36	0.130	.720	.004

Table A.3-13. Results of the Two-way GLM Repeated Measures Analysis on **VELOCITY**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,36	0.139	.711	.004
Posture *	1,36	7.192	.011	.167
Age group x Posture	1,36	1.424	.241	.038

Table A.3-14. Results of the Two-way GLM Repeated Measures Analysis on **ANGULAR DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,36	3.587	.066	.091
Posture ***	1,36	21.570	.000	.375
Age group x Posture	1,36	0.063	.804	.002

Table A.3-15. Results of the Two-way GLM Repeated Measures Analysis on **ECCENTRICITY**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,36	1.309	.260	.035
Posture **	1,36	8.102	.007	.184
Age group x Posture	1,36	0.309	.582	.009

Table A.3-16. Results of the Two-way GLM Repeated Measures Analysis on **LINE DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group *	1,36	6.091	.018	.145
Posture ***	1,36	27.977	.000	.428
Age group x Posture	1,36	0.351	.558	.010

Table A.3-17. Results of the Two-way GLM Repeated Measures Analysis on **ELLIPSE DEVIATION**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group **	1,36	11.486	.002	.242
Posture	1,36	2.337	.135	.061
Age group x Posture	1,36	0.149	.702	.004

Table A.3-18. **Age group** differences of gesture performance parameters.
 Results of the Mann-Whitney test. (p-values for exact two-tailed comparison).
 For dependent variables which are non-normal distributed in one or both of the age groups

Dependent variable	older		younger		<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Error rate *	14.91	10.93	8.11	5.95	111.50	-2.017	.044	-.33
Velocity	5.97	2.73	6.25	1.80	140.00	-1.182	.246	-.19
Ellipse deviation ***	0.45	0.11	0.35	0.06	71.00	-3.197	.001	-.52

Table A.3-19. **Posture** differences of gesture performance parameters.
 Results of the Wilcoxon test. (p-values for exact two-tailed comparison).
 For dependent variables which are non-normal distributed in one or both of the age groups

Dependent variable	table		hand		<i>T</i>	<i>z</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Error rate ***	9.21	10.18	13.82	9.92	110.50	-3.356	.000	-.54
Velocity ***	6.97	3.65	5.26	2.23	69.00	-4.372	.000	-.71
Ellipse deviation	0.38	0.10	0.41	0.13	244.00	-1.835	.067	-.30

A.4 Study 2A: A User-derived Gesture Set

A.4.1 Elements of the Gesture Description Language

Table A.4-1. Elements of the gesture description language; (* = denotes mandatory elements)

Abbreviation	Name	Description	Examples
sign *	sign	A gesture visually depicting a symbol (e.g., an alphanumeric character), or icon (e.g. an arrow).	<i>e.g. arrow, arrowhead, letter(s), number(s,) word, plus sign, minus sign, cross, check mark, scissors, question mark, magnifying glass, bracket...</i>
motion *	motion	The basic type of motion used in indexical and direct manipulation gestures.	<i>tap, double tap, triple tap, swipe, curve, pinch, spread, press & hold, circle, square, triangle, spiral, loop</i>
attr	attribute	Attribute of the sign or motion, used to describe more complicated spatial patterns.	<i>dashed, curved, angled, cursive, semi-, quarter-</i>
char	character	The sequence of one or more characters for alphanumeric gestures.	<i>e.g. 1, 2, a, A, SMS, yes, no, ok...</i>
dir	direction	The direction in which a sign or motion is performed.	<i>right, left, up, down, diagonal, diagonal up, diagonal down, clockwise, counter-clockwise, vertical, horizontal</i>
finger	number of fingers	Specifies the number of fingers or the order of finger with which a gesture part is performed.	<i>with two fingers, with three fingers, with 2nd finger, with 3rd finger</i>
rel _{space}	spatial relation	A local preposition describing the relation of the sign or motion to the screen or interface objects.	<i>in, on, at, to, across, around, above, below, towards, left of, right of</i>
loc	location	Specifies where on the screen or over which interaction objects the gesture is performed. Objects could be e.g., an <i>icon</i> , a <i>textbox</i> , a <i>list entry</i> , a <i>progress bar</i> , etc.	<i>object, multiple objects, another object, neutral space, the corner, the center, the top, the bottom, upper border, lower border, right border, left border</i>
rel _{temp}	temporal relation	The temporal relation in which different parts of the gesture are performed. The number of repetitions can also be coded.	<i>simultaneously, sequentially, twice, repeatedly</i>
sequence	sequence	A temporal preposition connecting different parts of a multipart gesture.	<i>then, while</i>

A.4.2 Analysis of Gesture Types

Table A.4-2. Frequencies of gesture types according to **age group**.
Results from the chi-square test. (p-values for exact two-tailed comparison;
* denotes significant effect at the Bonferroni-corrected of $\alpha = .0083$)

	older	younger			
gesture type	% of total gestures		df	χ^2	p
manipulative *	41.67	67.63	1	151.531	.000
indexical	19.96	18.08	1	1.285	.280
iconic *	24.45	8.73	1	100.068	.000
symbolic	2.02	1.41	1	1.208	.327
alphanumeric *	8.97	2.56	1	42.581	.000
combined	2.93	1.59	1	4.570	.044

A.4.3 Correlation Analysis

Table A.4-3. Correlation Table between **age**, **technology experience**, and **gesture characterization**. Each cell shows Spearman's rho (r_s) and p (two-tailed testing). N=53.

Parameters: AGE =age, HANDY = mobile phone experience, COMP=computer literacy, TOUCH = touchscreen experience, GEST = gesture interaction experience, MANI = frequency of manipulative gestures, SYM = frequency of symbolic gestures, INDEX = frequency of indexical gestures, ALPHA = frequency of alphanumeric gesture, COMBI = frequency of combined gestures, COMPLEX = gesture complexity index, FINGER = average number of fingers used

	AGE	HANDY	COMP	TOUCH	GEST	MANI	SYM	ICON	INDEX	ALPHA	COMBI	COMPLEX	FINGER
AGE	1 ,000	-,697** ,000	-,605** ,000	-,502** ,000	-,420** ,002	-,584** ,000	,146 ,296	,462** ,000	,029 ,837	,404** ,003	,271* ,049	,481** ,000	-,476** ,000
HANDY	-,697** ,000	1	,378** ,005	,266 ,054	,393** ,004	,279* ,043	-,282* ,041	-,216 ,120	,035 ,806	-,237 ,088	-,024 ,863	-,255 ,065	,336* ,014
COMP	-,605** ,000	,378** ,005	1	,468** ,000	,364** ,007	,302* ,028	-,050 ,721	-,188 ,178	-,096 ,496	-,220 ,113	-,167 ,232	-,447** ,001	,326* ,017
TOUCH	-,502** ,000	,266 ,054	,468** ,000	1	,336* ,014	,333* ,015	,023 ,873	-,163 ,244	-,309* ,024	-,174 ,211	,008 ,953	-,112 ,425	,414** ,002
GEST	-,420** ,002	,393** ,004	,364** ,007	,336* ,014	1	,276* ,045	-,231 ,096	-,133 ,344	-,109 ,436	-,197 ,157	-,170 ,223	-,151 ,280	,477** ,000
MANI	-,584** ,000	,279* ,043	,302* ,028	,333* ,015	,276* ,045	1	-,360** ,008	-,854** ,000	,144 ,305	-,726** ,000	-,581** ,000	-,718** ,000	,456** ,001
SYM	,146 ,296	-,282* ,041	-,050 ,721	,023 ,873	-,231 ,096	-,360** ,008	1	,327* ,017	-,324* ,018	,347* ,011	,067 ,634	,276* ,046	-,100 ,477
ICON	,462** ,000	-,216 ,120	-,188 ,178	-,163 ,244	-,133 ,344	-,854** ,000	,327* ,017	1	-,511** ,000	,506** ,000	,479** ,000	,705** ,000	-,399** ,003
INDEX	,029 ,837	,035 ,806	-,096 ,496	-,309* ,024	-,109 ,436	,144 ,305	-,324* ,018	-,511** ,000	1	-,324* ,018	-,296* ,032	-,371** ,006	,028 ,843
ALPHA	,404** ,003	-,237 ,088	-,220 ,113	-,174 ,211	-,197 ,157	-,726** ,000	,347* ,011	,506** ,000	-,324* ,018	1	,541** ,000	,606** ,000	-,293* ,033
COMBI	,271* ,049	-,024 ,863	-,167 ,232	,008 ,953	-,170 ,223	-,581** ,000	,067 ,634	,479** ,000	-,296* ,032	,541** ,000	1	,550** ,000	-,233 ,093
COMPLEX	,481** ,000	-,255 ,065	-,447** ,001	-,112 ,425	-,151 ,280	-,718** ,000	,276* ,046	,705** ,000	-,371** ,006	,606** ,000	,550** ,000	1	-,171 ,221
FINGER	-,476** ,000	,336* ,014	,326* ,017	,414** ,002	,477** ,000	,456** ,001	-,100 ,477	-,399** ,003	,028 ,843	-,293* ,033	-,233 ,093	-,171 ,221	1

Table A.4-4. Partial correlation between **age**, **technology experience**, and **gesture characterization**. Each cell shows correlation coefficient r and p (two-tailed testing). $df = 51$ (47).

Lower section (grey): controlled for COMP, HANDY, TOUCH and GEST

Parameters as in Table 8-28

	AGE	COMPLEX	FINGER	MANI	ICON	ALPHA	COMBI	COMP	HANDY	TOUCH	GEST
AGE	1,000 ,000	,481 ,000	-,476 ,000	-,584 ,000	,462 ,000	,404 ,003	,271 ,049	-,605 ,000	-,697 ,000	-,502 ,000	-,420 ,002
COMPLEX	,481 ,000	1,000	-,171 ,221	-,718 ,000	,705 ,000	,606 ,000	,550 ,000	-,447 ,001	-,255 ,065	-,112 ,425	-,151 ,280
FINGER	-,476 ,000	-,171 ,221	1,000	,456 ,001	-,399 ,003	-,293 ,033	-,233 ,093	,326 ,017	,336 ,014	,414 ,002	,477 ,000
MANI	-,584 ,000	-,718 ,000	,456 ,001	1,000	-,854 ,000	-,726 ,000	-,581 ,000	,302 ,028	,279 ,043	,333 ,015	,276 ,045
ICON	,462 ,000	,705 ,000	-,399 ,003	-,854 ,000	1,000	,506 ,000	,479 ,000	-,188 ,178	-,216 ,120	-,163 ,244	-,133 ,344
ALPHA	,404 ,003	,606 ,000	-,293 ,033	-,726 ,000	,506 ,000	1,000	,541 ,000	-,220 ,113	-,237 ,088	-,174 ,211	-,197 ,157
COMBI	,271 ,049	,550 ,000	-,233 ,093	-,581 ,000	,479 ,000	,541 ,000	1,000	-,167 ,232	-,024 ,863	,008 ,953	-,170 ,223
COMP	-,605 ,000	-,447 ,001	,326 ,017	,302 ,028	-,188 ,178	-,220 ,113	-,167 ,232	1,000	,378 ,005	,468 ,000	,364 ,007
HANDY	-,697 ,000	-,255 ,065	,336 ,014	,279 ,043	-,216 ,120	-,237 ,088	-,024 ,863	,378 ,005	1,000	,266 ,054	,393 ,004
TOUCH	-,502 ,000	-,112 ,425	,414 ,002	,333 ,015	-,163 ,244	-,174 ,211	,008 ,953	,468 ,000	,266 ,054	1,000	,336 ,014
GEST	-,420 ,002	-,151 ,280	,477 ,000	,276 ,045	-,133 ,344	-,197 ,157	-,170 ,223	,364 ,007	,393 ,004	,336 ,014	1,000
AGE	1,000 ,010	,365 ,010	-,207 ,153	-,508 ,000	,454 ,001	,310 ,030	,363 ,010				
COMPLEX	,365 ,010	1,000	-,059 ,689	-,739 ,000	,721 ,000	,595 ,000	,553 ,000				
FINGER	-,207 ,153	-,059 ,689	1,000	,312 ,029	-,350 ,014	-,184 ,206	-,211 ,145				
MANI	-,508 ,000	-,739 ,000	,312 ,029	1,000	-,857 ,000	-,701 ,000	-,615 ,000				
ICON	,454 ,001	,721 ,000	-,350 ,014	-,857 ,000	1,000	,469 ,001	,490 ,000				
ALPHA	,310 ,030	,595 ,000	-,184 ,206	-,701 ,000	,469 ,001	1,000	,546 ,000				
COMBI	,363 ,010	,553 ,000	-,211 ,145	-,615 ,000	,490 ,000	,546 ,000	1,000				

Table A.4-5. Partial correlation between **technology experience**, and **gesture characterization** with **age** controlled for. Each cell shows correlation coefficient r and p (two-tailed testing). $df=50$
Compare differences to table 8-29 (upper section); Parameters as in Table 8-28

	COMPLEX	FINGER	MANI	ICON	ALPHA	COMBI	COMP	HANDY	TOUCH	GEST
COMPLEX	1,000 	,075 ,597	-,614 ,000	,620 ,000	,513 ,000	,497 ,000	-,224 ,110	,127 ,369	,171 ,226	,064 ,653
FINGER	,075 ,597	1,000 	,250 ,074	-,230 ,101	-,125 ,378	-,123 ,386	,055 ,699	,006 ,964	,231 ,100	,347 ,012
MANI	-,614 ,000	,250 ,074	1,000 	-,811 ,000	-,660 ,000	-,541 ,000	-,079 ,576	-,220 ,117	,057 ,690	,042 ,765
ICON	,620 ,000	-,230 ,101	-,811 ,000	1,000 	,394 ,004	,414 ,002	,130 ,357	,167 ,237	,090 ,524	,076 ,590
ALPHA	,513 ,000	-,125 ,378	-,660 ,000	,394 ,004	1,000 	,490 ,000	,034 ,813	,068 ,630	,036 ,800	-,033 ,815
COMBI	,497 ,000	-,123 ,386	-,541 ,000	,414 ,002	,490 ,000	1,000 	-,004 ,980	,239 ,088	,174 ,218	-,064 ,650
COMP	-,224 ,110	,055 ,699	-,079 ,576	,130 ,357	,034 ,813	-,004 ,980	1,000 	-,077 ,587	,238 ,089	,152 ,281
HANDY	,127 ,369	,006 ,964	-,220 ,117	,167 ,237	,068 ,630	,239 ,088	-,077 ,587	1,000 	-,135 ,339	,154 ,274
TOUCH	,171 ,226	,231 ,100	,057 ,690	,090 ,524	,036 ,800	,174 ,218	,238 ,089	-,135 ,339	1,000 	,160 ,257
GEST	,064 ,653	,347 ,012	,042 ,765	,076 ,590	-,033 ,815	-,064 ,650	,152 ,281	,154 ,274	,160 ,257	1,000

A.5 Study 2B: Evaluations of the Gesture Set

A.5.1 Analysis of Type, Number of Fingers, and Complexity

Table A.5-1. Results of the Two-way GLM Analysis on **SUITABILITY RATING**;

* = significant at Bonferroni-corrected $\alpha = .0167$

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group *	1,40	16.090	.000	.287
Type *	1.54,61.49	43.031	.000	.518
Age group x Type	1.54,61.49	2.682	.090	.063

Table A.5-2. Results of the Two-way GLM Analysis on **SUITABILITY RATING**;

* = significant at Bonferroni-corrected $\alpha = .0167$

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group	1,40	0.061	.806	.002
Number of Fingers	1,40	2.220	.144	.053
Age group x Number of Fingers	1,40	10.500	.002	.208

Table A.5-3. Results of the Two-way GLM Analysis on **SUITABILITY RATING**;

* = significant at Bonferroni-corrected $\alpha = .0167$

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Age group *	1,40	14.960	.000	.272
Complexity *	1.83,73.09	48.185	.000	.546
Age group x Complexity	1.83,73.09	3.513	.039	.081

Table A.5-4. Results of the independent t-test on **SUITABILITY RATING** for each gesture type

Dependent variable	older		younger		<i>df</i>	<i>T</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Indexical gestures *	4.88	0.60	4.46	0.66	40	-2.164	.037	.32
Manipulative gestures	4.62	0.37	4.41	0.57	40	-1.439	.158	.22
Iconic gestures ***	3.73	0.82	2.75	0.75	40	-4.010	.000	.54
Symbolic gestures *	3.81	1.11	2.88	1.30	40	-2.498	.017	.37

A.5.2 Analysis of Individual Gestures

Table A.5-5. Results of the independent t-test on **SUITABILITY RATING** for each individual gesture

task	gesture	older		younger		df	T	p	r	
		M	SD	M	SD					
Select multiple (grid)	tap on multiple objects sequentially	5.15	1.66	5.09	1.85	40	.108	.914	0.02	
	swipe right, then down, then left across multiple objects	5.20	1.47	5.18	1.56	40	.039	.969	0.01	
	square around multiple objects	4.70	1.63	4.45	1.82	40	.459	.648	0.07	
	circle across multiple objects	5.00	1.45	4.59	1.74	40	.824	.415	0.13	
Select single	Tap on object	5.85	1.09	5.91	1.69	40	-.133	.895	0.02	
	Circle on Object	4.00	1.34	2.73	1.39	40	3.021	.004	0.43	**
	Double Tap on Object	5.20	1.51	5.32	1.76	40	-.233	.817	0.04	
	Checkmark on object	4.50	1.47	3.36	1.50	40	2.478	.018	0.36	*
Select multiple (list)	Swipe across multiple objects	6.00	0.86	5.73	1.42	40	.744	.461	0.12	
	Tap on multiple object sequentially	5.35	1.18	5.18	1.99	34.7	.336	.739	0.06	
	Bracket across multiple objects	3.35	1.46	2.00	1.27	40	3.201	.003	0.45	**
	Swipe diagonal down	4.40	1.64	4.00	1.45	40	.841	.405	0.13	
Move object	Curve up from object to another object	4.85	1.69	5.50	1.47	40	-1.330	.191	0.21	
	Swipe up from object to another object	5.20	1.01	6.05	1.00	40	-2.731	.009	0.40	**
	Arrow up from object	3.90	1.80	2.64	1.92	40	2.195	.034	0.33	*
	Tap on object, then Tap on another object	4.50	1.79	2.45	1.41	40	4.137	.000	0.55	***
Delete	Cross on object	4.75	1.71	4.41	2.11	40	.572	.571	0.09	
	Swipe right across object	5.25	1.55	4.59	1.76	40	1.280	.208	0.20	
	Swipe right, then swipe left on object repeatedly	3.70	2.03	2.91	1.95	40	1.288	.205	0.20	
	Tap on Object, then swipe right	4.90	1.12	4.45	1.68	36.8	1.018	.315	0.17	
Confirm	Tap on object	5.30	1.49	5.14	1.83	40	.315	.754	0.05	
	Checkmark on object	4.80	1.44	5.32	1.25	40	-1.250	.218	0.19	
	Double tap on object	5.00	1.17	5.23	1.15	40	-.634	.530	0.10	
	Word 'ja' on object	3.45	2.09	1.77	1.15	28.9	3.178	.004	0.51	**
Cancel	Cross on object	3.85	1.31	4.36	1.89	40	-1.013	.317	0.16	
	Tap on neutral space	2.65	1.50	4.45	2.18	37.4	-3.155	.003	0.46	**
	Letter N on object	3.75	2.07	3.36	2.06	40	.605	.549	0.10	
	Tap on object	3.05	2.04	2.45	1.71	40	1.029	.310	0.16	
Open	Double Tap on object	5.35	1.42	6.18	0.96	40	-2.239	.031	0.33	*
	Tap on object	5.65	1.46	5.82	1.47	40	-.372	.712	0.06	

	Checkmark	4.05	1.36	2.95	1.17	40	2.805	.008	0.41	**
	Arrow down from object	3.30	1.22	2.59	1.47	40	1.693	.098	0.26	
Close	Tap on object	5.55	1.36	5.86	1.04	40	-.846	.402	0.13	
	Double tap on object	5.50	1.05	5.77	1.19	40	-.783	.438	0.12	
	Cross on object	4.30	1.53	2.86	1.42	40	3.156	.003	0.45	**
	Half-Circle counter-clockwise on object	3.65	1.39	3.64	1.40	40	.032	.975	0.01	
Scroll down	Swipe up across multiple objects	4.05	2.04	5.14	2.21	40	-1.651	.107	0.25	
	Swipe down across multiple objects	5.20	1.67	4.18	2.08	40	1.734	.091	0.26	
	Arrow up across multiple objects	3.15	1.53	2.59	1.22	40	1.314	.196	0.20	
	Swipe down with two fingers across multiple objects	4.50	1.82	3.64	1.97	40	1.473	.149	0.23	
Scroll up	Swipe up across multiple objects	5.40	1.31	4.09	2.35	33.6	2.255	.031	0.36	*
	Swipe down across multiple objects	4.00	2.18	5.18	2.13	40	-1.777	.083	0.27	
	Arrow down across multiple objects	2.74	1.33	2.18	1.05	39	1.493	.144	0.23	
	Swipe down with two fingers across multiple objects	3.21	1.96	4.09	2.02	39	-1.410	.166	0.22	
Previous page	Arrow left	3.65	1.66	2.68	1.70	40	1.862	.070	0.28	
	Swipe left	5.55	1.47	4.45	1.74	40	2.195	.034	0.33	*
	Swipe left on object	5.05	1.67	5.23	1.85	40	-.325	.747	0.05	
	Swipe right	2.35	1.14	3.36	2.08	33.1	-1.981	.056	0.33	
Next page	Arrow right	3.85	1.57	2.50	1.60	40	2.763	.009	0.40	**
	Swipe left	2.70	1.87	3.23	2.14	40	-.848	.402	0.13	
	Swipe right	5.20	1.40	4.59	1.76	40	1.231	.225	0.19	
	Swipe right on object	5.00	1.65	5.27	1.80	40	-.509	.614	0.08	
Zoom in	Spread diagonal	4.00	2.20	5.67	1.77	36.5	-2.664	.011	0.40	*
	Double Tap	4.30	1.38	4.14	1.86	40	.321	.750	0.05	
	Circle clockwise	4.55	1.67	3.82	1.82	40	1.355	.183	0.21	
	Tap	5.00	1.65	3.33	1.62	39	3.256	.002	0.46	**
Pan	Swipe left	5.35	1.50	6.45	0.80	40	-3.021	.004	0.43	**
	Swipe right	3.85	2.28	2.68	1.81	40	1.849	.072	0.28	
	Arrow left	3.20	1.28	2.36	1.18	40	2.205	.033	0.33	*
	Swipe left with two fingers	4.30	1.59	4.55	2.02	40	-.435	.666	0.07	
Rotate	Quarter-circle clockwise	5.10	1.25	4.27	1.75	40	1.745	.089	0.27	
	Curved arrow clockwise	3.40	1.43	2.05	1.17	40	3.369	.002	0.47	**
	Press and hold while quarter-circle clockwise with	4.30	1.75	5.14	2.08	40	-1.404	.168	0.22	
	3/4 Circle clockwise	4.90	1.21	4.05	1.84	36.6	1.795	.081	0.28	
Zoom out	Pinch diagonal	4.15	1.95	4.77	2.58	40	-.875	.387	0.14	
	Press and hold	4.45	1.67	3.05	1.62	40	2.768	.009	0.40	**

	Tap	4.65	1.69	2.50	1.82	40	3.952	.000	0.53	***
	Spread diagonal	3.05	2.14	3.55	2.54	40	-.680	.500	0.11	
Select text	Swipe right across multiple objects	5.32	1.34	5.36	1.81	39	-.095	.925	0.02	
	Curve diagonal down from object to object	2.15	0.81	2.50	1.79	29.9	-.827	.415	0.15	
	Swipe left across multiple objects	4.30	0.98	3.64	1.62	35.0	1.623	.114	0.26	
	Tap on multiple objects twice	4.95	1.47	3.95	1.50	40	2.173	.036	0.32	*
Open menu	Tap on object	4.50	1.79	3.09	1.69	40	2.624	.012	0.38	*
	Press and hold on object	4.55	1.67	5.50	1.65	40	-1.850	.072	0.28	
	Letter M on object	3.55	1.93	2.32	1.59	40	2.267	.029	0.34	*
	Swipe right across object towards right border	4.50	1.47	3.73	2.00	40	1.413	.165	0.22	
Help	Question mark across multiple objects	4.25	1.55	3.41	2.22	40	1.410	.166	0.22	
	Tap on object	5.55	1.50	4.36	2.08	38.1	2.130	.040	0.33	*
	Letter H on neutral space	3.60	1.67	2.91	1.90	40	1.247	.220	0.19	
	Tap with two fingers simultaneously	3.00	1.41	2.55	1.41	40	1.044	.303	0.16	
Play	Double Tap on object	5.10	1.29	5.64	1.43	40	-1.269	.212	0.20	
	Tap on object	5.55	1.19	5.73	1.12	40	-.497	.622	0.08	
	Swipe right on object	4.85	0.93	3.36	1.71	33.1	3.545	.001	0.52	**
	Checkmark on object	4.45	1.79	2.59	1.10	30.9	4.008	.000	0.58	***
Pause	Tap on object	5.40	1.23	5.82	1.47	40	-.995	.326	0.16	
	Tap on neutral space	5.15	1.53	5.14	2.01	40	.025	.981	0.00	
	Pause sign	3.55	1.76	2.64	1.65	40	1.736	.090	0.26	
	Press and hold on object	4.85	1.50	4.05	1.25	40	1.895	.065	0.29	
Stop	Tap on object	4.95	1.36	4.82	1.84	40	.262	.795	0.04	
	Double tap on object	4.60	1.31	4.64	1.50	40	-.083	.934	0.01	
	Swipe left on object	4.90	1.52	2.95	1.91	40	3.624	.001	0.50	**
	Press and hold on object	4.85	1.09	4.41	1.99	33.1	.901	.374	0.15	
Fast forward	Swipe right on object	5.20	1.64	6.23	0.87	28.3	-2.498	.019	0.43	*
	Tap on object	5.05	1.10	5.73	1.20	40	-1.899	.065	0.29	
	Swipe right on neutral space	4.85	1.53	3.91	1.57	40	1.962	.057	0.30	
	Arrowhead right twice	2.75	1.25	1.41	0.67	28.3	4.273	.000	0.63	***
Rewind	Swipe left on object	5.30	1.42	6.09	0.81	40	-2.245	.030	0.33	*
	Tap on object	5.20	1.20	5.77	1.27	40	-1.500	.141	0.23	
	Swipe left on neutral space	4.85	1.46	3.82	1.30	40	2.426	.020	0.36	*
	Arrowhead left twice	2.65	1.35	1.45	0.74	28.8	3.514	.001	0.55	**
next title	Swipe right on object	5.32	1.20	3.27	1.61	39	4.541	.000	0.59	***
	Number of the title	4.74	1.82	4.05	2.06	39	1.131	.265	0.18	
	Arrow right	4.45	1.39	3.41	2.09	40	1.881	.067	0.29	
	Swipe left on neutral space	2.20	1.15	3.68	2.59	29.6	-2.433	.021	0.41	*

Previous title	Number of the title	4.35	2.03	3.95	1.89	40	.653	.517	0.10	
	Arrow left	4.47	1.50	3.45	2.22	37.1	1.740	.090	0.27	
	Swipe left on object	4.63	1.42	2.23	1.45	39	5.351	.000	0.65	***
	Swipe right on neutral Space	2.30	1.30	3.64	2.57	31.7	-2.152	.039	0.36	*
Volume up	Swipe up from object	5.25	1.33	5.18	1.68	40	.145	.886	0.02	
	Tap on object, then Swipe up	5.00	1.49	5.50	1.22	40	-1.194	.240	0.19	
	Swipe up	4.75	1.52	3.68	1.73	40	2.119	.040	0.32	*
	Plus sign above object	2.35	1.57	1.73	1.12	40	1.493	.143	0.23	
Volume down	Tap on object, then swipe down	4.55	1.47	5.50	1.30	40	-2.224	.032	0.33	*
	Swipe down to object	4.90	1.21	3.68	1.99	35.2	2.425	.021	0.38	*
	Swipe down from object	4.75	1.55	3.86	2.05	40	1.565	.125	0.24	
	Tap on object	3.55	1.67	2.32	1.13	40	2.824	.007	0.41	**
Take call	Tap on object	5.55	1.19	5.68	1.59	40	-.302	.764	0.05	
	Tap	5.40	1.60	4.82	2.34	37.3	.946	.350	0.15	
	Check mark	3.60	1.60	2.91	1.57	40	1.410	.166	0.22	
	Double Tap on object	5.25	1.21	4.09	1.57	40	2.660	.011	0.39	*
End call	Cross on neutral space	3.95	2.01	2.82	1.92	40	1.866	.069	0.28	
	Tap	5.25	1.33	4.00	2.20	35.0	2.247	.031	0.35	*
	Tap on object	5.35	1.53	4.91	1.77	40	.859	.395	0.13	
	Double Tap	4.90	1.12	4.27	1.67	36.9	1.443	.157	0.23	
Make call	Double Tap on object	5.40	1.14	5.50	1.22	40	-.273	.786	0.04	
	Tap on object	6.05	0.89	4.82	1.87	30.6	2.769	.009	0.45	**
	Swipe right across object	4.26	1.48	3.14	1.58	39	2.339	.025	0.35	*
	Cross on object	3.21	1.44	1.55	0.74	39	4.762	.000	0.61	***
Write message	Tap on object	4.55	2.11	3.91	1.90	40	1.035	.307	0.16	
	Swipe right across object	4.75	1.62	4.05	1.81	40	1.324	.193	0.20	
	Letter SMS on object	3.20	1.77	1.64	1.40	40	3.196	.003	0.45	**
	Letter N on object	3.40	1.43	2.50	1.74	40	1.822	.076	0.28	
Send message	Checkmark on object	4.58	1.54	3.91	1.82	39	1.260	.215	0.20	
	Double tap on object	5.15	0.99	4.14	2.17	30.0	1.979	.057	0.34	
	Double tap in object	5.20	1.20	3.95	2.13	33.7	2.366	.024	0.38	*
	Tap on object	5.47	1.39	3.59	1.89	39	3.579	.001	0.50	**

A.6 Study 3: Gesture Interaction on a Functional Prototype

A.6.1 Interaction Steps

Table A.6-1. Minimal number of interaction steps for each task and scenario (A & B).

Interaction steps were calculated conservatively, such that e.g. one key press was assumed for each letter to delete, or one swipe gesture to advance the side scrolling list by one picture. However, a long press on the delete button enabled to continuously delete consecutive letters (similar to a long press on the backspace of the PC keyboard), or a quickly executed swipe allowed for momentum-scrolling several items of the list at once.

condition	type of interaction	calling		message		calendar		gallery		web	
		A	B	A	B	A	B	A	B	A	B
key	keystroke	10	12	56	59	27	27	19	26	14	16
touch & gesture	tap	4	4	8	8	5	5	8	8	5	5
	direct manipulation gesture	0	0	1	1	1	1	1	6	0	0
	symbolic gesture	2	2	3	3	2	2	3	3	4	4

A.6.2 Analysis of Input Mode and Order of Interaction

Table A.6-2. Results of the Two-way GLM Repeated Measures Analysis on **TIME ON TASK**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Order	1,30	0.010	.920	.000
Input mode **	1,30	15.071	.001	.334
Order x Input mode	1,30	3.326	.078	.100

Table A.6-3. Results of the Two-way GLM Repeated Measures Analysis on **REDUNDANT INTERACTION STEPS**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Order	1,30	.398	.533	.013
Input mode ***	1,30	55.697	.000	.650
Order x Input mode	1,30	0.054	.818	.002

Table A.6-4. Results of the Two-way GLM Repeated Measures Analysis on **PERCENTAGE INTERACTION STEPS**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Order	1,30	0.517	.478	.017
Input mode	1,30	0.237	.630	.008
Order x Input mode	1,30	0.021	.884	.001

Table A.6-5. Results of the Two-way GLM Repeated Measures Analysis on **REDUNDANT SCREEN CHANGES**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Order *	1,30	5.910	.021	.165
Input mode ***	1,30	55.513	.000	.649
Order x Input mode	1,30	.043	.836	.001

A.6.3 Analysis of Input Mode and Task

Table A.6-6. Results of the Two-way GLM Repeated Measures Analysis on **TIME ON TASK**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Input mode **	1,31	14.019	.001	.311
Task ***	4,124	19.858	.000	.390
Input mode x Task *	2.94,91.06	4.031	.010	.115

Table A.6-7. Results of the Two-way GLM Repeated Measures Analysis on **REDUNDANT INTERACTION STEPS**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Input mode ***	1,31	126.642	.000	.803
Task ***	1.79,55.53	41.84	.000	.574
Input mode x Task ***	1.69,52.39	30.729	.000	.598

Table A.6-8. Results of the Two-way GLM Repeated Measures Analysis on **PERCENTAGE INTERACTION STEPS**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Input mode	1,31	0.010	.820	.000
Task ***	4,124	11.734	.000	.275
Input mode x Task *	3.04,94.17	3.774	.013	.109

Table A.6-9. Results of the Two-way GLM Repeated Measures Analysis on **REDUNDANT SCEEN CHANGES**

Effect	<i>df</i>	<i>F</i>	<i>p</i>	<i>part.η²</i>
Input mode ***	1,31	57.280	.000	.649
Task ***	4,124	5.432	.000	.149
Input mode x Task **	3.10,96.03	5.252	.002	.145

A.6.4 Analysis of Questionnaire Data

Table A.6-10. Comparison of **gesture mode** and **key mode** on **SUMI** and **QUESI** subscales. Results of the Wilcoxon signed-rank test. (p-values for exact two-tailed comparison).

Test	Subscale	gestures		keys		<i>T</i>	<i>z</i>	<i>p</i>	<i>r</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
SUMI	Usability (global) ***	55.81	12.13	45.63	14.16	51.50	-3.451	.000	-.43
	Efficiency *	48.84	13.55	42.56	13.78	118.50	-2.142	.031	-.20
	Affect ***	56.90	11.90	45.09	15.38	62.50	-3.354	.000	-.42
	Helpfulness ***	58.58	9.63	44.50	14.64	9.50	-4.408	.000	-.55
	Controllability *	53.52	11.85	47.56	11.32	122.0	-2.470	.012	-.31
	Learnability **	48.68	15.57	40.28	16.77	82.50	-2.920	.003	-.37
QUESI	Intuitive Use (global) *	2.95	0.82	2.60	0.75	123.0	-2.450	.013	-.31
	Perc. mental workload	2.25	1.01	2.50	0.85	137.0	-1.256	.215	-.16
	Perc. achievement * of goals	3.72	0.76	3.26	0.78	77.50	-2.513	.011	-.31
	Perc. effort of learning*	2.19	0.98	2.54	0.91	98.50	-2.388	.015	-.30
	Familiarity	2.58	0.91	2.30	0.81	136.0	-1.772	.077	-.22
	Perc. error rate	2.13	0.97	2.55	0.96	92.00	-1.917	.055	-.24

A.6.5 Analysis of Prior Experience

Table A.6-10. Correlation table of **experience** parameters (vertical) and **interaction performance** parameters (horizontal). r_s = Spearman's rho, r = Pearson's correlation coefficient.

Parameters: HANDY = mobile phone experience, TOUCH = touchscreen experience, GEST = gesture interaction experience, COMP = computer literacy, TIME = time on task, STEPS = redundant interaction steps, STEP% = percentage interaction steps, SCREEN = redundant screen changes

		gesture mode				key mode			
		TIME	STEPS	STEP %	SCREEN	TIME	STEPS	STEP %	SCREEN
HANDY	r_s	-,211	,087	,079	-,130	-,250	-,180	-,172	-,197
	p	,247	,637	,668	,477	,168	,323	,348	,281
TOUCH	r_s	-,194	-,028	-,021	-,080	-,134	-,079	-,052	,010
	p	,288	,880	,908	,662	,464	,669	,778	,957
GEST	r_s	-,365 [*]	-,156	-,171	-,308	-,087	-,063	-,058	-,080
	p	,040	,395	,351	,087	,637	,733	,753	,663
COMP	r	-,435 [*]	-,456 ^{**}	-,470 ^{**}	-,187	-,277	-,087	-,068	-,243
	p	,013	,009	,007	,305	,125	,637	,713	,181

Table A.6-11. Results of the Multivariate Analysis of Covariance (MANCOVA). IV=Input mode, DV = time-on- task, redundant interaction steps, percentage interaction steps, redundant screen changes, COVARIATES = mobile phone experience, touchscreen experience, gesture experience, computer literacy; V = Pillai's trace

Effect	V	df	F	p	part. η^2
Input mode ***	.861	4,24	37.117	.000	.861
Mobile experience *	.360	4,24	3.371	.025	.360
Computer literacy	.284	4,24	2.376	.080	.284
Gesture experience	.133	4,24	0.920	.469	.133
Touch experience	.105	4,24	0.707	.595	.105
Input mode x Mobile experience	.114	4,24	0.771	.555	.114
Input mode x Computer literacy	.306	4,24	2.650	.058	.306
Input mode x Gesture experience	.166	4,24	1.196	.339	.166
Input mode x Touch experience	.049	4,24	0.306	.871	.049

A.7 Excursus: A Brief History of Multi-Touch Technology

The development of 2D gestural interfaces is intrinsically related to the technological development of touch sensitive surfaces. It was not until 2006-2008 that multi-touch technology received public attention, and along with it, new ways of technology interaction were introduced. In 2006, Jeff Han gave a demonstration at the TED conference which, as Saffer put it, “will likely go down in computer history near the ‘Mother of All Demos’ presentation that Doug Engelbart made in 1968” (Saffer, 2008, p. 1). Han demonstrated on a table-sized screen with amazing perfection, speed and fluidity, how objects on the screen could be selected and dragged around, how multiple finger input could be used to simultaneously rotate and zoom screens, to pan, tilt and scale items effortlessly, which has previously been only known from fictional work, for example the well-known gesture interaction scenes from the 2002 science fiction movie *Minority Report*. One year later, this technology was commercially shipped in wall-sized dimensions. Also in 2007, the first iPhone was released, and brought multi-touch technology on small screen devices to the mass consumer market. Again one year later, in 2008, Microsoft launched its *Microsoft Surface* interactive table, which combined multi-touch technology with the ability to sense objects and the idea of a collaborative, multi-user interaction. Both the Microsoft Surface Table and Han’s wall-sized multi-touch display were heavily used in the US television coverage of the US presidential elections of 2008, and thus gained wide public attention. However, this recent technological breakthrough was preceded by decades of research on multi-touch technology, which can be retraced vividly in Bill Buxton’s Blog “Multi-Touch Systems that I Have Known and Loved” (Buxton, 2011), which is briefly summarized in the following. In Buxton’s view, the origin of multi-touch research can be traced back to regular keyboards, which were designed for intentional multi-key presses (such as for example CTR + ALT + DEL) as well as unintentional simultaneous key presses (e.g. when in fast typing, one pushes the next key before releasing the old one, the so-called *n-key rollover*). The earliest concepts of touch-screen interfaces date back to the mid-sixties, e.g. from Johnson (1965, 1967). The development of regular (“single-touch”) screens started in the late sixties, at IBM, and a range of universities independently, and by 1971 several techniques had been presented. One of the first systems to be generally known was the touch screen terminal for the Plato IV computer which was released in 1972. The display developed by Herot & Weinzapfel in 1978 at MIT is remarkable for sensing eight different signals (e.g. pressure and torque) from a single touch point, especially since modern touchscreen devices mostly rely only on position (Herot & Weinzapfel, 1978). The “Flexible Machine Interface” by Mehta could probably be regarded as the first multi-touch system designed for human input to a computer system (Mehta, 1982): behind a frosted-glass panel a camera was installed which could record black spots where

finger presses occurred, and even finger pressure could be derived from the size of the black spots. Even though the *video place* system presented by Myron Krueger in 1983 isn't a multi-touch device in a strict sense since it relied on dwell-time instead of touch to trigger events, it was way ahead of its time and hugely influenced all following gesture-interaction systems (e.g. Krueger, Gionfriddo, & Hinrichsen, 1985). The user could interact with the system using multiple fingers, both hands, and a rich variety of gestures. For example, the now popular pinching and spreading gesture to scale objects was already incorporated in this system. The first multi-touch *screen*, meaning that multi-touch was incorporated not in a separate input device such as a pad, panel or tablet, but was overlaid directly on the primary computer screen (a CRT, at that time), was developed by Boie and colleagues at the Bell Labs in 1984. The system used a capacitive array of transparent touch sensors, which allowed for very fast manipulation of graphical objects with the fingers. The late eighties saw several prototypes of multi-touch type interfaces, e.g. a multi-touch tablet by Bill Buxton and his group at the University of Toronto, or the flip keyboard – a multi-touch pad integrated in the bottom of a keyboard – which he developed later at Xerox PARC. Already in 1992 IBM and Bell South released the *Simon*, a touchscreen-based device that “was arguably the world's first smart phone” (Buxton, 2011). While the screen of this mobile phone was only single touch, it is remarkable that already 15 years before the iPhone was released, this device already passed on any physical buttons. Thanks to the flexibility that the touchscreen interface offered, it included already applications such as calendar, address book, calculator, note pad, sketch pad, email and games. Even though the technology was already available, it still found applications only in research facilities, or in niche products, such as the multi-touch tablets and keyboards produced by FingerWorks since 1998 (later purchased by Apple), or the *Lemur*, a music controller featuring a transparent multi-touch surface developed by Jazz Mutant in 2003. A nice overview of gesture controlled user interfaces, comprising both research prototypes and commercial products from the past thirty years, can also be found in Bhuiyan & Picking, (2009). This brief historic account of the development of multi-touch technology shows that it took more than four decades from the first concepts and prototypes to a technology sufficiently matured to be included in mass consumer products. As of today, where multi-touch technology is commonly available in various devices, the time is ripe to investigate the full potential of this technology with respect to a more senior friendly interaction.

A.8 Study Materials

The following section contains the materials that were used by the experimenter or provided to the participants in the empirical studies. The standardized questionnaires that were used are not listed in the following, but can be obtained by their respective authors.

description	used in study	page
Structured interview on participant demographics and previous experience with mobile phones. The interview was conducted at the start of the session.	1A and 1B	262
Questionnaire assessing Mental and Physical Demand (NASA-TLX), subjective difficulty of selected gestures, prior and future use of touch and gesture.	1A	263
Questionnaire on NASA-TLX Mental Demand and Physical demand and operation posture preference. Administered separately for 1-finger and 2-finger gestures. Questionnaire on prior and future use of touch and gesture	1B	264
Structured interview on participant demographics	2A, 2B and 3	265
Questionnaire assessing prior experience with mobile phones, touchscreens and gesture control	2A and 2B	266
Questionnaire on the general criteria of gesture evaluation; Section of the participants' evaluation sheet for each task and gesture	2B	267
Instructions for using the functional prototype in gesture mode	3	268
Instructions for using the functional prototype in key mode	3	269
Task instructions for study 3 (scenario A)	3	270
Task instructions for study 3 (scenario B)	3	271
Questionnaire on gesture recall and functionality	3	272
Questionnaire assessing prior experience with mobile phones, touchscreens and gesture control	3	273

Strukturierte Befragung zur Person	
Vp-Nr:	
1. Wie alt sind Sie? _____ Jahre	
2. Geschlecht: ☐ weiblich ☐ männlich	
3. Sind Sie Rechts- oder Linkshänder? ☐ Rechtshänder ☐ Linkshänder	
4. Welchen höchsten Schulabschluss besitzen Sie? ☐ Hauptschule ☐ Realschule ☐ Abitur / Fachabitur ☐ Fachhochschule / Universität ☐ Sonstiges: _____	
5. Was ist Ihre derzeitige Tätigkeit? ☐ berufstätig ☐ arbeitslos ☐ Student ☐ Rentner / pensioniert ☐ Sonstiges: _____	
6. Wie ist Ihre derzeitige Wohnsituation? ☐ alleine ☐ mit Partner ☐ bei Familie ☐ betreut ☐ andere: _____	
7. Leiden Sie an einer Sehbeeinträchtigung, die nicht mit einer Brille oder Kontaktlinsen ausgeglichen werden kann? ☐ ja ☐ nein Wenn ja, wie äußerst sich die Beeinträchtigung: _____ _____ _____	

8. Leiden Sie unter Erkrankungen, die Ihnen die Bewegung von Armen und /oder Händen erschweren? ☐ ja ☐ nein Wenn ja, wie äußerst sich die Beeinträchtigung: _____ _____ _____	
9. Benutzen Sie ein Mobiltelefon? ☐ ja ☐ nein Wenn ja, seit wann: _____	
10. Wie häufig schreiben Sie Kurznachrichten mit dem Mobiltelefon? ☐ nie ☐ ca. _____ Mal pro ☐ Tag ☐ Woche ☐ Monat	
11. Spielen Sie ein Musikinstrument? ☐ ja, und zwar: _____ ☐ nein	
12. Haben Sie ein anders Hobby bei dem Sie Ihre Hand/Finger besonders stark einsetzen, wie beispielsweise Zeichnen, Malen, Stricken, Häkeln etc...? ☐ ja, und zwar: _____ ☐ nein	

Fragebogen zur Aufgabe
Vp-Nr: _____

1. Beanspruchung durch die Aufgabe:

Markieren Sie bitte auf den folgenden Skalen, in welchem Maße Sie sich in den beiden genannten Dimensionen von der **Aufgabe die Gesten schnell und genau auszuführen** beansprucht oder gefordert gesehen haben:

Geistige Anforderungen

Wie viel geistige Anstrengung war bei der Informationsaufnahme und -verarbeitung erforderlich (z.B. Denken, Entscheiden, Rechnen, Erinnern, Hinsehen, Suchen...)? War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex, erforderte sie hohe Genauigkeit oder war sie fehlertolerant?

Körperliche Anforderungen

Wie viel körperliche Aktivität war erforderlich (z.B. Ziehen, Drücken, Drehen, Steuern, Aktivieren,...)? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholend oder mühselig?

2. Bitte beurteilen Sie die Schwierigkeit der gezeigten Gesten. Begutachten Sie die Gesten auf der Tafel und versuchen sich zu erinnern welche davon Ihnen in der Ausführung besonders leicht oder besonders schwer vorgekommen sind. Wählen Sie bitte 5 besonders einfache und 5 besonders schwierige Gesten aus (im Vergleich zu den anderen) und schreiben deren Nummern in nachstehende Liste.

Besonders einfache Gesten: _____ Besonders schwierige Gesten: _____

Nr.: _____ Nr.: _____

Nr.: _____ Nr.: _____

Nr.: _____ Nr.: _____

Nr.: _____ Nr.: _____

Nr.: _____ Nr.: _____

3. Was haben Sie an den „schwierigeren“ Gesten als besonders schwierig empfunden?

4. Haben Sie schon einmal ein technisches Gerät über einen berührungsempfindlichen Bildschirm bedient (z.B. Fahrkartenautomaten)?

☐ nein, noch nie

☐ ja, und zwar: _____

5. Haben Sie schon einmal ein technisches Gerät mit ähnlichen Gesten bedient?

☐ nein, noch nie

☐ ja, und zwar: _____

6. Können Sie sich vorstellen technische Geräte mit Hilfe der getesteten oder ähnlicher Gesten zu bedienen?

☐ trifft gar nicht zu ☐ trifft eher wenig zu ☐ trifft eher stark zu ☐ trifft sehr stark zu

7. Falls ja, bei welchen technischen Geräten bzw. Anwendungen erscheint Ihnen solche Bedienung sinnvoll?

Fragebogen zur Aufgabe
Vp-Nr: _____

1. Beanspruchung durch die Aufgabe:

Markieren Sie bitte auf den folgenden Skalen, in welchem Maße Sie sich in den beiden genannten Dimensionen von der **Aufgabe die Gesten schnell und genau auszuführen** beansprucht oder gefordert gesehen haben:

Geistige Anforderungen

Wie viel geistige Anstrengung war bei der Informationsaufnahme und -verarbeitung erforderlich (z.B. Denken, Entscheiden, Rechnen, Erinnern, Hinsehen, Suchen...)?
War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex, erforderte sie hohe Genauigkeit oder war sie fehlertolerant?

Körperliche Anforderungen

Wie viel körperliche Aktivität war erforderlich (z.B. Ziehen, Drücken, Drehen, Steuern, Aktivieren,...)? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholend oder mühselig?

2. Welche der drei Haltungen fanden Sie am angenehmsten?

Tisch ☐ Linke Hand ☐ Rechte Hand ☐

3. Was haben Sie an den „schwierigeren“ Gesten als besonders schwierig empfunden?

Abschlussfragebogen
Vp-Nr: _____

1. Haben Sie schon einmal ein technisches Gerät über einen berührungsempfindlichen Bildschirm bedient (z.B. Fahrkartenautomaten)?

☐ nein, noch nie

☐ ja, und zwar: _____

2. Haben Sie schon einmal ein technisches Gerät mit ähnlichen Gesten bedient?

☐ nein, noch nie

☐ ja, und zwar: _____

3. Können Sie sich vorstellen technische Geräte mit Hilfe der getesteten oder ähnlicher Gesten zu bedienen?

☐ trifft gar nicht zu ☐ trifft eher wenig zu ☐ trifft eher stark zu ☐ trifft sehr stark zu

4. Falls ja, bei welchen technischen Geräten bzw. Anwendungen erscheint Ihnen solche Bedienung sinnvoll?

Befragung zur Person
Vp-Nr: _____

1. Wie alt sind Sie?
_____ Jahre

2. Geschlecht

☐ weiblich ☐ männlich

3. Sind Sie Rechts- oder Linkshänder?

☐ rechts ☐ links

4. Welchen höchsten Schulabschluss besitzen Sie?

☐ Hauptschule ☐ Abitur/ Fachabitur

☐ Realschule ☐ Fachhochschule / Universität

☐ sonstiges: _____

5. Was ist Ihre derzeitige Tätigkeit?

☐ berufstätig ☐ arbeitslos

☐ Rentner/ pensioniert ☐ Student

☐ sonstiges: _____

6. Leiden Sie an einer Sehbeeinträchtigung, die nicht mit einer Brille oder Kontaktlinsen ausgeglichen werden kann?

☐ ja ☐ nein

Wenn ja, wie äußert sich die Beeinträchtigung? _____

7. Leiden Sie unter Erkrankungen, die Ihnen die Bewegung von Armen und/ oder Händen erschweren?

☐ ja ☐ nein

Wenn ja, wie äußert sich die Beeinträchtigung? _____

Abschlussfragebogen

Vp-Nr: _____

1. Benutzen Sie ein Mobiltelefon?

☐ ja ☐ nein

Wenn ja, seit wie vielen Jahren ? _____ Jahre

2. Wie oft nutzen Sie folgende Funktionen bei Ihrem Mobiltelefon?

	nie	selten	gelegentlich	oft
Telefonieren	☐	☐	☐	☐
Kurznachrichten (SMS)	☐	☐	☐	☐
Uhrzeit	☐	☐	☐	☐
Wecker	☐	☐	☐	☐
Kalender, Termine	☐	☐	☐	☐
Kamerafunktion	☐	☐	☐	☐
Musik / Radiofunktion	☐	☐	☐	☐
Internet	☐	☐	☐	☐
Spiele	☐	☐	☐	☐
Sonstiges: _____	☐	☐	☐	☐

3. Wie viele Kurznachrichten schreiben Sie im Durchschnitt mit dem Handy?

☐ ca. _____ pro ... ☐ Tag ☐ Woche ☐ Monat ☐ Jahr

☐ keine

4. Haben Sie schon einmal ein technisches Gerät über einen *berührungs-empfindlichen Bildschirm* (Touchscreen) bedient?

☐ Geld-/ Bankautomat ☐ Infoterminal ☐ Fahrkartenautomaten

☐ Mobiltelefon ☐ Digitalkamera ☐ Navigationsgerät

☐ MP3 Spieler ☐ Sonstiges: _____ ☐ nein, noch nie

5. Haben Sie schon einmal ein technisches Gerät mit *Fingergesten* bedient die ähnlich zu den Gesten waren die in dem heutigen Versuch abgefragt wurden?

☐ Nein, noch nie ☐ ja, und zwar folgende Geräte: _____

6. Wie stark sind Sie mit dem Bedienkonzept des Produktes „iPhone“ bzw. „iPod Touch“ vertraut?

☐ Ich kenne das Produkt ☐ Ich habe von dem Produkt gehört, weiß aber nichts über die Bedienung. ☐ Durch Beobachtung oder Werbung kenne ich die Grundzüge der Bedienung.

☐ Ich habe die Bedienung selbst schon einmal ausprobiert. ☐ Ich habe das Gerät selbst schon öfters bedient. ☐ Ich besitze eines der genannten Produkte.

7. „Ich würde Geräte, die über *Fingergesten* bedient werden, in Zukunft selbst gerne benutzen“.

☐ stimmt völlig ☐ stimmt ziemlich ☐ stimmt teils ☐ stimmt wenig ☐ stimmt gar nicht

8. Bei welchen technischen Geräten oder Anwendungen scheint Ihnen eine Bedienung durch *Fingergesten* geeignet?

Wie geeignet erscheinen Ihnen die gezeigten Vorschläge?

Aktion	Video	gar nicht geeignet	nicht geeignet	wenig geeignet	mittel- mäßig geeignet	ziemlich geeignet	sehr geeignet	außer- ordentlich geeignet	Video	Aktion
1	A	☐	☐	☐	☐	☐	☐	☐	A	1
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
2	A	☐	☐	☐	☐	☐	☐	☐	A	2
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
3	A	☐	☐	☐	☐	☐	☐	☐	A	3
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
4	A	☐	☐	☐	☐	☐	☐	☐	A	4
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
5	A	☐	☐	☐	☐	☐	☐	☐	A	5
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
6	A	☐	☐	☐	☐	☐	☐	☐	A	6
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
7	A	☐	☐	☐	☐	☐	☐	☐	A	7
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
8	A	☐	☐	☐	☐	☐	☐	☐	A	8
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
9	A	☐	☐	☐	☐	☐	☐	☐	A	9
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
10	A	☐	☐	☐	☐	☐	☐	☐	A	10
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	
11	A	☐	☐	☐	☐	☐	☐	☐	A	11
	B	☐	☐	☐	☐	☐	☐	☐	B	
	C	☐	☐	☐	☐	☐	☐	☐	C	
	D	☐	☐	☐	☐	☐	☐	☐	D	

Fragebogen Gestenbewertung

Vp-Nr: _____

Welche Merkmale waren Ihnen *generell* bei der Auswahl der Gesten wichtig?

	nicht wichtig	wenig wichtig	mittel- mäßig wichtig	ziemlich wichtig	sehr wichtig
Die Geste kann mit nur einem Finger ausgeführt werden.	☐	☐	☐	☐	☐
Die Geste kann besonders schnell ausgeführt werden.	☐	☐	☐	☐	☐
Die Geste ist unverwechselbar.	☐	☐	☐	☐	☐
Die Geste kann gut gemerkt werden.	☐	☐	☐	☐	☐
Die Geste stellt ein vertrautes Zeichen dar.	☐	☐	☐	☐	☐
Die Geste besteht aus möglichst wenigen Strichen.	☐	☐	☐	☐	☐
Die Geste ist intuitiv verständlich.	☐	☐	☐	☐	☐
Die Geste ist mühelos mit dem Finger zu zeichnen.	☐	☐	☐	☐	☐
Die Geste kann auch bei einhändiger Bedienung des Gerätes ausgeführt werden.	☐	☐	☐	☐	☐
Die Geste würde selbst bei ungenauer Ausführung funktionieren.	☐	☐	☐	☐	☐
Die Geste ist plausibel nachvollziehbar.	☐	☐	☐	☐	☐
Die Geste enthält keine Rundungen.	☐	☐	☐	☐	☐
Sonstiges Merkmal: _____	☐	☐	☐	☐	☐

Instruktion Gestensteuerung

Vorabinformation:

Der folgende Bildschirm wird als „Startbildschirm“ bezeichnet. Von diesem Bildschirm aus beginnen Sie jede Aufgabe und sollen nach Beendigung einer Aufgabe wieder dorthin zurück kehren.

Sie sollen das Gerät nicht über Tasten, sondern lediglich mit Hilfe von Fingerbewegungen auf dem Bildschirm bedienen. Einige Elemente lassen sich durch leichten Fingerdruck aktivieren, andere Funktionen lassen sich nur mit Hilfe von Bewegungsmustern, so genannten „Gesten“, aktivieren.

Eine Beispielgeste ist das **Fragezeichen**:

Diese Geste funktioniert überall und mit deren Hilfe können Sie jederzeit eine Hilfeseite aufrufen, die Ihnen darüber Auskunft gibt, welche weiteren Gesten momentan für Aktionen zur Verfügung stehen. Die Hilfeseite sieht z.B. so aus:

Auf den Bildschirmen sind jeweils unterschiedliche Gesten möglich. Das heißt, manche Gesten funktionieren nur in bestimmten Bildschirmen, andere Gesten auf anderen Bildschirmen. Um zu sehen, welche Gesten möglich sind, können Sie jederzeit die Hilfeseite mit der Geste **Fragezeichen** aufrufen.

Gesten funktionieren fast überall, nur bei Fenstern im **Vordergrund**, wie folgenden Beispiele zeigen, sind sie nicht möglich:

Bitte drücken Sie ruhig etwas kräftiger mit Ihrem Finger auf, Sie können nichts kaputt machen!

Bei Schaltflächen wie dieser hier

sollten Sie versuchen Ihre
Fingerspitze mittig zu platzieren.

Versuchen Sie bitte auch das Gerät **nur über den Bildschirm** zu bedienen, und keine der Tasten auf dem Gerät zu drücken!

Haben Sie noch weitere Fragen? Dann stellen Sie diese bitte jetzt dem Versuchleiter. Ansonsten können Sie nun mit einem kurzen Training beginnen. Viel Spaß!

Instruktion Tastensteuerung

Der folgende Bildschirm wird als „Startbildschirm“ bezeichnet. Von diesem Bildschirm aus beginnen Sie jede Aufgabe und sollen nach Beendigung einer Aufgabe wieder dorthin zurück kehren.

Sie sollen das Gerät **nicht über den Bildschirm**, sondern lediglich mit Hilfe der Tasten unterhalb des Bildschirms bedienen. Folgende Tasten finden dabei Verwendung:

Bitte beachten Sie dass in manchen Situationen die mittlere OK-Taste **länger gedrückt** werden muss um bestimmte Funktionen aufzurufen.

Haben Sie noch weitere Fragen? Dann stellen Sie diese bitte jetzt dem Versuchleiter. Viel Spaß!

4

Szenario A:

1

Stellen Sie sich vor, Ihr Enkel Tobias kommt am Samstag zu Besuch. Sie wollen Ihre Tochter Susanne anrufen, um mit ihr zu klären, um welche Uhrzeit Sie Ihren Enkel vom Bahnhof abholen sollen.

- 1.1. Rufen Sie Ihre Tochter **Susanne** Schneider an. Ihre Nummer ist im Gerät gespeichert.
- 1.2. Kehren Sie danach zum **Startbildschirm** zurück.

2

Ihre Tochter Susanne hebt nicht ab, anscheinend ist sie gerade nicht zu erreichen. Also versuchen Sie es mit einer Nachricht (SMS).

- 2.1 Schreiben Sie Ihrer Tochter **Susanne** eine Nachricht (SMS). Der Text wird automatisch eingefügt. Schicken Sie die Nachricht bitte noch nicht ab.

Ihnen gefällt der gerade getippte Text nicht mehr, Sie möchten einen Satz daraus löschen. Danach ist der Text in Ordnung und Sie können die Nachricht versenden.

- 2.2 Markieren Sie den Satz „**Ich frage wegen Tobias**“.
- 2.3 Löschen Sie den markierten Satz.
- 2.4 Schicken Sie die Nachricht ab.

3

Sie werden nun von Ihrer Tochter zurückgerufen. Ihr Enkel kommt schon um 10:00 Uhr am Bahnhof an. Das ist toll, so können Sie ihn schon eher sehen! Sie dachten, er würde um 15:00 Uhr kommen. Gestern hatten Sie allerdings in Ihren Kalender eingetragen, dass Tobias um 15:00 Uhr kommt, und möchten den Kalendereintrag nun verschieben.

- 3.1 Gehen Sie zum Kalendereintrag für **Samstag, 5. März**
- 3.2 Verschieben Sie den Eintrag von 15:00 Uhr auf **10:00 Uhr**.
- 3.3 Kehren Sie auf den **Startbildschirm** zurück.

Sie machen Sich Gedanken was Sie mit Ihrem Enkel unternehmen möchten. Sie fassen den Plan mit ihm auf den Berliner Fernsehturm zu steigen. Vom Fernsehturm haben Sie vor kurzem ein Foto gemacht. Jetzt möchten Sie das Foto Ihrer Tochter Susanne schicken, damit Sie es Ihrem Enkel Tobias zeigen kann.

- 4.1 Suchen Sie zunächst dieses Foto vom **Berliner Fernsehturm**:

- 4.2 Schicken Sie dann das Foto an **Susanne**.

5

Sie möchten einen kleinen Spaziergang machen um die Gegend rund um den Fernsehturm noch etwas zu erkunden. Vielleicht gibt es noch weitere interessante Dinge für Tobias in der Nähe. Da Sie nicht wissen wie das Wetter morgen ist möchten Sie im Internet auf einer ihrer Lieblingsseiten „Wetter.de“ nachsehen...

- 5.1 Wählen Sie das Internet
- 5.2 Lassen Sie sich die **Favoritenliste** anzeigen.
- 5.3 Wählen Sie die Seite **wetter.de** aus.
- 5.4 Finden Sie heraus welche Temperatur für **morgen Mittag** vorhergesagt wird.
- 5.5 Öffnen Sie anschließend noch Ihre Favoritenliste „**web.de**“ und lesen Sie die Schlagzeile vor.
- 5.6 Kehren Sie danach auf den **Startbildschirm** zurück.

Szenario B:

1

Ihr Freund Otto wollte neulich mit Ihnen Boule spielen, doch leider mussten Sie ihm absagen. Diese Woche hätten Sie Zeit. Sie versuchen Ihren Freund Otto anzurufen, um sich mit ihm zum Boule spielen zu verabreden.

- 1.1 Rufen Sie Ihren Freund **Otto** an. Seine Nummer ist im Gerät gespeichert.
- 1.2 Kehren Sie danach zum **Startbildschirm** zurück.

2

Ihr Freund Otto hebt nicht ab, anscheinend ist er gerade nicht zu erreichen. Also versuchen Sie es mit einer Nachricht (SMS).

- 2.1 Schreiben Sie Ihrem Freund **Otto** eine Nachricht (SMS). Der Text wird automatisch eingefügt. Schicken Sie die Nachricht bitte zunächst noch nicht ab.

Ihnen gefällt der gerade getippte Text nicht mehr, Sie möchten einen Satz daraus löschen. Danach ist der Text in Ordnung und Sie können die Nachricht versenden.

- 2.2 Markieren Sie den Satz „**Das wäre doch toll**“.
- 2.3 Löschen Sie den markierten Satz.
- 2.4 Schicken Sie die Nachricht ab.

3

Ihr Freund Otto ruft zurück und teilt Ihnen mit, dass er am 15ten März Zeit hat und gerne mit Ihnen Boule spielen würde. Sie verabreden sich für 11 Uhr und sie tragen diesen Termin auf dem Kalender in Ihrem Handy ein. Kurze Zeit später ruft Otto zurück dass es ihm am Nachmittag besser passt, und Sie einigen sich darauf den Termin auf 15 Uhr zu verschieben.

- 3.1 Gehen Sie zum Kalendereintrag für **Dienstag, 15. März**
- 3.2 Verschieben Sie den Eintrag von 11:00 Uhr auf **15:00 Uhr**.
- 3.3 Kehren Sie auf den **Startbildschirm** zurück.

4

Im Gespräch mit Otto kommt auch seine bevorstehende Reise nach München zur Sprache. Sie kennen sich gut aus in München und geben ihm Tipps was er tun könnte. Bei schönem Wetter empfiehlt sich ein Spaziergang im „Englischen Garten“. Da er sich darunter nichts vorstellen kann möchten Sie ihm ein Bild vom Englischen Garten in München schicken, welches Sie kürzlich mit ihrem Handy aufgenommen haben.

- 4.1 Suchen Sie zunächst ein Foto vom **Englischen Garten** in **München**:

- 4.2 Schicken Sie dann das Foto an **Otto**.

5

Da der „Englische Garten“ jedoch nur bei gutem Wetter einen Besuch lohnt, wollen Sie sich nun noch mit Hilfe Ihres Handys über die Wetteraussichten informieren. Übermorgen wird Otto in München sein, und es wäre hilfreich zu wissen, ob Regen vorausgesagt wird. Sie wollen diese Informationen im Internet suchen ...

- 5.1 Wählen Sie das Internet
- 5.2 Lassen Sie sich die **Favoritenliste** anzeigen.
- 5.3 Wählen Sie die Seite **wetter.de** aus.
- 5.4 Finden Sie heraus ob für **übermorgen Mittag** Regen vorhergesagt ist.
- 5.5 Öffnen Sie anschließend noch Ihre Favoritenliste „**Spiegel online**“ und lesen Sie die Schlagzeile vor.
- 5.6 Kehren Sie danach auf den **Startbildschirm** zurück.

Fragebogen Handynutzung

Vp-Nr: _____

1. Benutzen Sie ein Mobiltelefon?

☐ ja☐ nein

Wenn ja, seit wie vielen Jahren ? _____ Jahre

Hersteller/Marke/Typ ihres aktuellen Mobiltelefons: _____

Wie bedienen Sie Ihr Gerät? ☐ Tasten ☐ Bildschirm ☐ anders: _____

2. Wie oft nutzen Sie folgende Funktionen bei Ihrem Mobiltelefon?

	nie	selten	gelegentlich	oft
Telefonieren	☐	☐	☐	☐
Kurznachrichten (SMS)	☐	☐	☐	☐
Uhrzeit	☐	☐	☐	☐
Wecker	☐	☐	☐	☐
Kalender, Termine	☐	☐	☐	☐
Kamerafunktion	☐	☐	☐	☐
Musik / Radiofunktion	☐	☐	☐	☐
Internet	☐	☐	☐	☐
Spiele	☐	☐	☐	☐
Bilder / Videos	☐	☐	☐	☐
Navigation / Karten	☐	☐	☐	☐
Sonstiges: _____	☐	☐	☐	☐

3. Wie viele Kurznachrichten schreiben Sie im Durchschnitt mit dem Handy?

☐ ca. _____ pro ... ☐ Tag ☐ Woche ☐ Monat ☐ Jahr☐ keine

Bitte wenden!

4. Haben Sie schon einmal ein technisches Gerät über einen *berührungsempfindlichen Bildschirm* (Touchscreen) bedient?

- ☐ Geld-/ Bankautomat ☐ Infoterminal ☐ Fahrkartenautomaten
- ☐ Mobiltelefon ☐ Digitalkamera ☐ Navigationsgerät
- ☐ MP3 Spieler ☐ Sonstiges: _____ ☐ nein, noch nie

5. Haben Sie schon einmal ein technisches Gerät mit *Fingergesten* bedient die ähnlich zu den Gesten waren die in dem heutigen Versuch verwendet wurden?☐ Nein, noch nie ☐ ja, und zwar folgende Geräte: _____

6. Wie stark sind Sie mit dem Bedienkonzept des Produktes „iPhone“ bzw. „iPod Touch“ vertraut?

- ☐ Ich kenne das Produkt ☐ Ich habe von dem Produkt gehört, weiß aber nichts über die Bedienung. ☐ Durch Beobachtung oder Werbung kenne ich die Grundzüge der Bedienung.
- ☐ Ich habe die Bedienung selbst schon einmal ausprobiert. ☐ Ich habe das Gerät selbst schon öfters bedient. ☐ Ich besitze eines der genannten Produkte.

7. „Ich würde Geräte, die durch *Fingergesten* bedient werden, in Zukunft selbst gerne benutzen“.

☐ stimmt völlig ☐ stimmt ziemlich ☐ stimmt teils ☐ stimmt wenig ☐ stimmt gar nicht

8. Bei welchen technischen Geräten oder Anwendungen scheint Ihnen eine Bedienung durch Fingergesten geeignet?

☐ bei keinem Gerät

☐ folgenden Geräten: _____