UNIVERSITY OF CAMBRIDGE
Department of Engineering

Exploring a Capability-Demand Interaction Model for Inclusive Design Evaluation

Dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy

Umesh Persad
St. Edmund's College
September 2011

Declaration

This thesis is the result of the author's work and does not comprise the outcome of work done in collaboration with others, unless clearly stated otherwise. References to work developed by other researchers is explicitly indicated in the text. This thesis has not been submitted in whole or in part for consideration for any other degree or qualification at this University or any other institute of learning. This thesis contains fewer than 65,000 words, including appendices, bibliography, footnotes and tables. This dissertation does not contain more than 150 figures.

Umesh Persad
St. Edmund's College
University of Cambridge

September 2011

Abstract

Designers are required to evaluate their designs against the needs and capabilities of their target user groups in order to achieve successful, inclusive products. This dissertation presents exploratory research into the specific problem of supporting analytical design evaluation for Inclusive Design. The analytical evaluation process involves evaluating products with user data rather than testing with actual users. The work focuses on the exploration of a capability-demand model of product interaction as the basis for analytical inclusive evaluation. This model suggests that by comparing the measured sensory, cognitive and motor capabilities of a user population to the corresponding product demands, the degree of fit between users and products can be assessed.

The research problem was addressed by firstly examining theories of human function and performance together with existing sources of user capability data. It was found that user capability data was fragmented and lacking in terms of predicting design exclusion and difficulty at the population level. More fundamentally, however, it was found that the relationships between measured capability in populations with low functional capacity and real world task performance with products (such as errors, times and difficulty) were not well understood. Given that an understanding of these relationships are necessary to guide capability data collection and to drive valid and robust analytical evaluation methods, the research effort focused on exploring these relationships via empirical and analytical studies.

The research process culminated in an experimental study with nineteen users of various functional capability profiles performing tasks with four consumer products (a clock radio, a mobile phone, a blender and a vacuum cleaner). Measures of user capability were related to corresponding product demands (on those capabilities) and task outcome measures. A complex picture emerged, where linear relationships did not generally account for significant variance in task outcome measures. Further, it appeared that multiple capabilities were possibly interacting in unknown ways to support real world interaction. These indicative results point to the further investigation of multivariate and non-linear models for describing capability-demand relationships, and also the replication of similar studies with larger sample sizes to confirm the relationships observed. The resulting overall recommendation, therefore, is that there is a need to direct research efforts in this critical but largely unexplored area of capability-demand model building for Inclusive Design evaluation.

Acknowledgements

I would like to thank my supervisors Prof. John Clarkson and Dr. Patrick Langdon for their guidance and support throughout the research process. I would also like to express thanks to all my colleagues at Cambridge Engineering Design Centre and members of the Inclusive Design and Healthcare research groups including James, Thomas, Suresh, Dawn, Carlos, Tim, Joy and Nick for their assistance and stimulating conversations. Many thanks to all the researchers in the extended iDesign research consortium for their periodical support, and the experts at Cambridge University and Addenbrookes Hospital who granted me interviews and access to whatever resources I required. Thanks to CAMSIGHT, the University of the Third Age (U3A) and the Papworth Trust Hester Adrian Centre for help in recruiting study participants who participated in user studies.

I would like to acknowledge St. Edmund's College for granting me a Commonwealth Studentship and the University of Trinidad and Tobago (UTT) for funding my final year of research. Additionally, I would like to thank my parents, grandparents, and in-laws for all their support in making my studies at Cambridge possible, and lots of love to my two daughters Vaishnavi and Vindi. Finally, my deepest gratitude to my wife Aarti for keeping me motivated. I know firsthand the struggles that you face in daily life, and hopefully the small contribution I have made with this research work will help deepen our understanding of how to design a more inclusive world.

Table of Contents

List of Figures

List of Tables

Chapter 1 **Introduction**

The research problem investigated in this thesis is introduced within the context of current challenges in Inclusive Design. Following from this, the main objectives and research questions are presented. Finally, a brief summary of each chapter is provided as a map through the thesis.

1.1 Research Motivation

Inclusive Design is defined as the "design of mainstream products and/or services that are accessible to, and usable by, people with the widest range of abilities within the widest range of situations without the need for special adaptation or design" (BSI, 2005). It therefore embodies a design philosophy that aims to consider the needs of people with reduced functional capacity in the design of products and services. The goal of Inclusive Design is to design products that are accessible to and usable by the maximum number of users without being stigmatising or resorting to special aids and adaptations (Keates & Clarkson, 2003a). This results in products that minimise the exclusion of less capable populations (Clarkson & Keates, 2003a).

The importance of Inclusive Design can be highlighted in the contexts of population ageing, legal considerations such as the Disability Discrimination Act in the UK, and the sociological and personal implications of independent living (Hosking, Waller, & Clarkson, 2010; Keates & Clarkson, 2003a). In addition, there are business drivers for Inclusive Design including untapped markets, opportunities for innovation, competitive advantage and brand recognition for accessible and easy to use products (Dong, 2004; Hosking et al., 2010; Keates & Clarkson, 2003a). The argument has been put forward that Inclusive Design is about *good* design (Coleman, 2006; Keates & Clarkson, 2003a), and ideally it should not be a sub-speciality, but rather it should be standard design practice.

1.1.1 *Population Ageing and Disability*

A population is considered to be ageing when the proportion of older persons (60 years and over) increases while the proportion of children (15 years and under) and the proportion of persons in the working ages (15 to 59) decreases (United Nations, 2007). In 2009, there were 737 million persons aged 60 years or over worldwide. This number is projected to increase to 2 billion in 2050 when older persons will outnumber children (United Nations, 2009). One out of every nine persons in the world is aged 60 or over, and by 2050 it is estimated that one person out of every five will be aged 60 or over (United Nations, 2009). Figure 1-1 and Figure 1-2 illustrate this striking and unprecedented demographic shift of the ageing of the world population from 2009 to 2050 respectively.

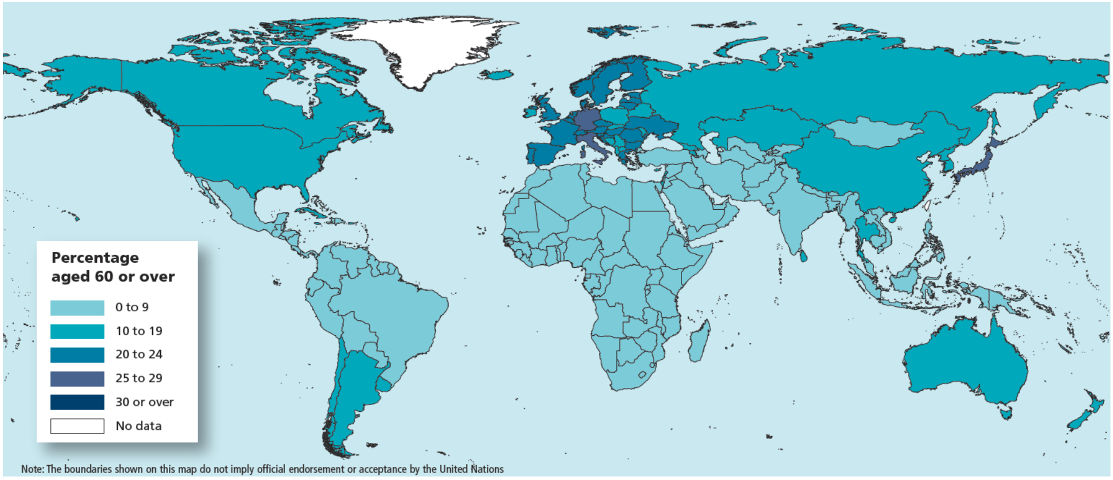


Figure 1-1 Percentage of the total population aged 60 years and over 2009 (United Nations, 2009)

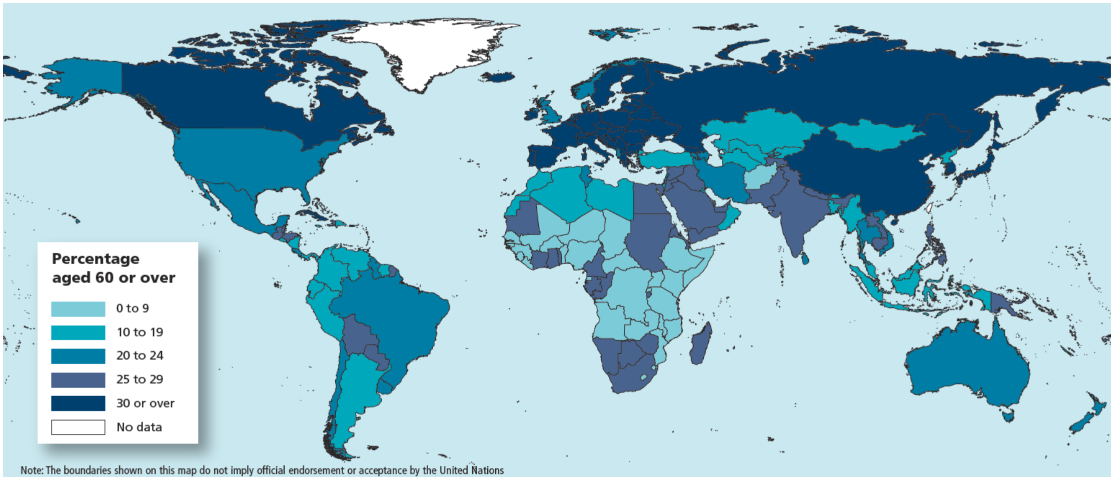


Figure 1-2 Percentage of the total population aged 60 years and over 2050 (United Nations, 2009)

This population ageing is pervasive in that it affects nearly all countries of the world. This occurrence is due to a reduction in fertility rates i.e. the growth of the number of children in the world is decreasing. Coupled with this, there is increased growth in the older population due to decreasing mortality rates (United Nations, 2002; United Nations, 2007; United Nations, 2009). In fact, the difference in growth rates between the older population and the total population is increasing as shown in Figure 1-3. In addition, the figure shows that the oldest-old population is the fastest growing group. Currently the oldest old (persons aged 80 years or over) constitute 14% of the population aged 60 or over, and by 2050, 20% of the older population are estimated to be 80 years or over (United Nations, 2009).

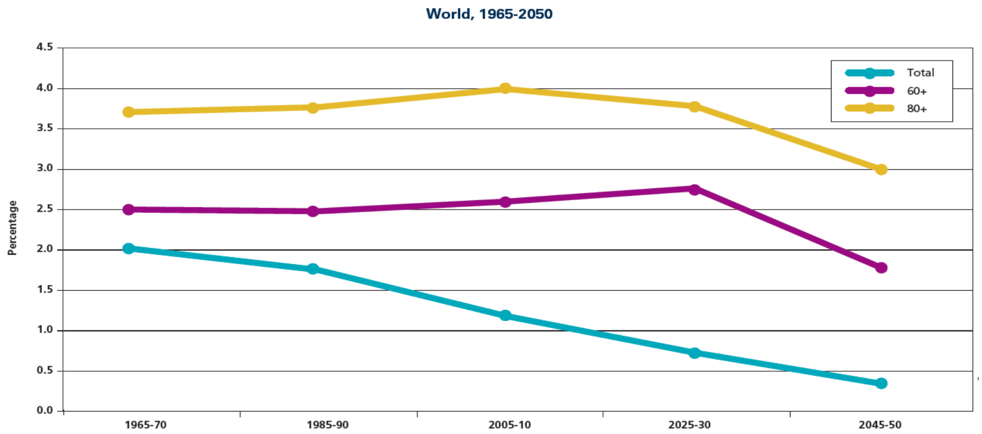


Figure 1-3 Average annual growth rate of total population, aged 60 or over and aged 80 or over (United Nations, 2009)

In the case of the United Kingdom, the population median age is expected to rise from 39.7 years in 2010 to 42.2 years by 2035. In keeping with the worldwide trend, the oldest age groups are expected to increase the fastest as shown in Figure 1-4. In 2010, there were 1.4 million people aged 85 and over. This number is projected to increase to 3.5 million by 2035, more than doubling over 25 years (Office for National Statistics, 2011).

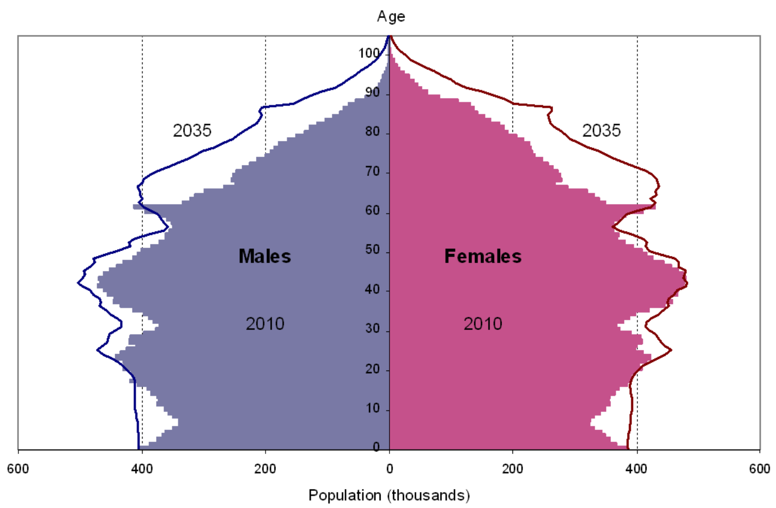


Figure 1-4 Estimated and projected age structure of the United Kingdom population, mid-2010 and mid-2035 (Office for National Statistics, 2011)

The global ageing phenomenon directly influences the trends of disability prevalence as there is a higher risk of disability in the older population (World Health Organisation, 2011). Based on 2010 global population estimates, more than a billion people (about 15% of the world's population) are estimated to live with some form of disability. The World Health Survey estimates that approximately 785 million (15.6%) persons 15 years and older live with a disability, while the Global Burden of Disease puts the figure at 975 million (19.4%) persons (World Health Organisation, 2011).

As the population ages, the number of people with disabilities will grow due to the increase in the number of persons with chronic health conditions, for example diabetes, cardiovascular diseases, and mental illness (World Health Organisation, 2011). This is clearly illustrated in Figure 1-5 from data across 59 countries. There is a disproportionately higher representation of older people evident in disability populations. It is estimated that there are over 10 million disabled people in Great Britain, comprising 0.8 million children, 5.1 million adults of working age and 5 million over state pension age (Office for Disability Issues, 2011).

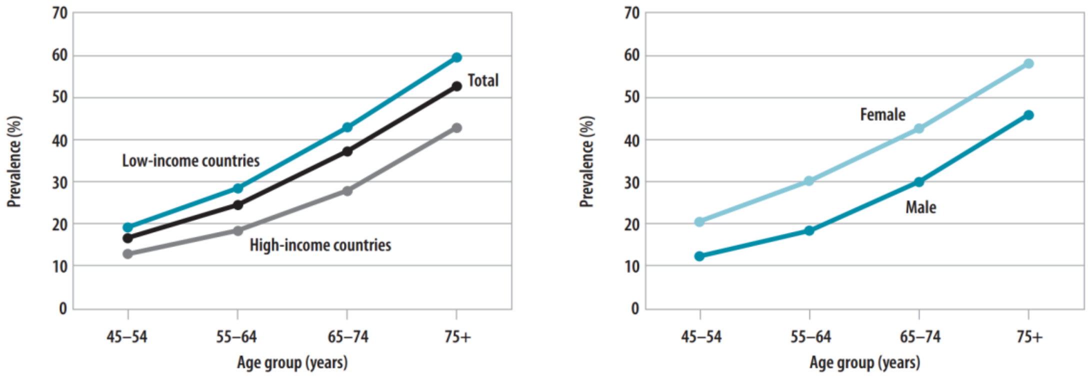


Figure 1-5 Age-specific disability prevalence, derived from multidomain functioning levels in 59 countries, by country income level and sex (World Health Organisation, 2011)

These trends in population ageing and disability have major implications for all aspects of life including politics, economic development, labour, family composition, housing, epidemiology and healthcare. Specifically, Inclusive Design becomes a necessary approach in designing environments, products and services that cater to the needs of an ageing population with various capability limitations. This approach is critical in order to remove barriers to access and participation in society, and also for enabling independent living for as long as possible.

1.1.2 Inclusive Product Design

In a review of Inclusive Design developments in the last 15 years, Coleman (Coleman, 2006) puts forward a threefold argument for Inclusive Design. Firstly, Inclusive Design grew out of a need to address social issues such as population ageing, disability and independent living via a design approach. Secondly, in order to effectively address these issues, design had to become *people* and *population* aware. Thirdly, this implies that business practice had to change in order to accommodate this mainstream approach to Inclusive Design. Various inclusively designed products have been brought to market including the Oxo Goodgrips line of kitchen products, BT Big Button Phones, the B&Q Sandbug, Fiskars Softouch scissors, the Ford Focus and Toyota Porte (Clarkson & Coleman, 2010; Coleman, 2006; Hosking et al., 2010; Mueller, 2003; Warburton, 2005). Though these products testify to the value of an inclusive approach to product development, they still remain few in number compared to the vast number of manufactured products on the market. The fact remains that inclusively designed products remain the exception, not the norm (Bontoft & Pullin, 2003).

Some progress has also been made in encouraging the uptake of Inclusive Design in industry via a multifaceted information dissemination strategy. This includes the provision of Inclusive Design Standards (BSI, 2005), books (Clarkson, Coleman, Keates, & Lebbon, 2003a; Preiser & Ostroff, 2001), design tools (Cardoso, 2003; Cardoso, 2005; Porter, Case, Marshall, Gyi, & Oliver, 2004), user data (Clarkson, Dong, & Keates, 2003b; Smith, Norris, & Peebles, 2000; Steenbekkers & VanBeijsterveldt, 1998) and various workshops (Coleman, 2006; Hosking et al., 2010). The use of the statement 'it is normal to be different' has had an impact on industry by moving the focus from disability to one of diversity over the whole population of potential customers (Hosking et al., 2010).

However, many key research questions remain (Coleman, 2006; Johnson, Clarkson, & Huppert, 2010). Is our understanding of human capability sufficient to provide designers with the user information they require? What user information is required to allow designers to make predictions of the numbers of people excluded by their design? Are existing data adequate and how should it be presented for use in the design process? In sum, there remains a need for a better understanding of how human capability data can support the inclusive approach (Johnson et al., 2010).

The research presented in this thesis specifically addresses the issue of using user data to support Inclusive Design evaluation. It is known that designers require supporting data, methods and tools in order to evaluate their designs (Clarkson & Keates, 2003b), and current user data is fragmented and lacking (Gyi, Sims, Porter, Marshall, & Case, 2004; Johnson et al., 2010; Persad & Clarkson, 2005). Previous research has also shown that quantitative data on the numbers of people with functional capability loss can be useful for designers as well as business managers and decision makers (Dong, 2004). Therefore, there is a need for user capability data that could enable the evaluation of design concepts throughout the design process.

In addition, the focus is placed on supporting an *analytical* evaluation framework. Analytical methods require the designer to analyse and inspect a given design without resorting to actual user trials. Analytical methods are especially advantageous in the Inclusive Design process where a population view on user capability is required (Carlsson, Iwarsson, & Sthål, 2002; Persad, Langdon, & Clarkson, 2005) and where time, cost and logistical constraints make testing with real users difficult (Goodman-Deane, Langdon, & Clarkson, 2010; Gyi et al., 2004). However, this does not negate the use of user trials and empirical studies, but rather aims to provide a supplemental method that addresses some of the deficiencies of empirical methods. This research also builds on previous research on the use of collected data for the evaluation of product designs (Cardoso, 2003; Clarkson et al., 2003b; Porter et al., 2004; Waller, Langdon, & Clarkson, 2010a).

1.2 Research Aim

The research presented in this thesis seeks to address the broad problem of supporting the evaluation of consumer products for Inclusive Design. The aim is to investigate the theoretical foundations of an evaluation framework that utilises measures of user capability to analytically evaluate consumer product designs. This *analytical* approach aims to be *predictive* of the real-world problems encountered by users of various levels of capability.

The approach investigated is grounded in the theoretical constructs of *user capabilities* and *product demands*. This entails understanding the interaction relationship between users' sensory, cognitive and motor capabilities and the consumer product's features. The relationships between user capabilities, product demands and task outcomes in disabled populations are not well understood (Steenbekkers & VanBeijsterveldt, 1998). If these relationships could be adequately modelled, the resulting models could be utilised as a valid and robust predictive tool in analytical product evaluation.

1.3 Research Questions

The following three exploratory research questions are investigated in this dissertation:

1. *What theoretical models exist for understanding the relationship between human functional capabilities and real world task performance in ageing/disabled populations?*
2. *What are the key elements of human functional capability that influence inclusive product interaction?*
3. *What relationships exist between measures of human functional capability and measures of task performance in the context of a user capability-product demand model of interaction?*

1.4 Scope

The research presented in this thesis is limited to investigating a capability-demand model of interaction as an approximation to real-world human interaction. Though the approach is envisioned to be useful to designers, this research seeks only to investigate the viability of the model as a basis for an analytical evaluation framework as a necessary first step. Thus the research questions are limited to understanding and investigating the model as a scientific basis for evaluation.

The definition of consumer products used in this thesis includes devices used in daily living, for activities such as cooking, cleaning, entertainment and communication. Though some of these products may contain interactive software interfaces, this work does not specifically address computer software products running on personal computers.

1.5 Thesis Outline

The following is an outline of the contents of the dissertation:

Chapter 2: Background

This chapter reviews relevant background literature in three areas: (1) the principles of Inclusive Design and the needs of designers, (2) current state-of-the-art tools and methods for evaluating products for Inclusive Design and (3) understanding the need for user data. Based on this review, gaps in knowledge will be highlighted with respect to analytical evaluation methods that can predict user problems for Inclusive Design.

Chapter 3: Research Approach

Chapter 3 describes the methodological approach employed to investigate the research questions. A three phase methodology is presented incorporating a mixture of qualitative and quantitative methods consisting of literature reviews, expert consultation, analytical studies, user observation and experimental studies.

Chapter 4: Theoretical Considerations

Models of human function and performance are reviewed and analysed, with the aim of understanding the relationship between low-level sensory, cognitive and motor functions and high-level user actions in performing real world tasks. A generic framework for the inclusive analytical evaluation of consumer products is developed.

Chapter 5: Capabilities and Demands

A review of the main elements of human sensory, cognitive and motor capability is presented based on existing literature and consultations with domain experts. This review distils the important underlying human capabilities that most impact users' interaction with consumer products. In addition, secondary data analysis of the 1996/97 Great Britain Disability Follow-up Survey (DFS) data is presented to explore the prevalence and co-occurrence of disability and health conditions in the UK population.

Chapter 6: Exploring Inclusive Interaction

Real world interaction is investigated in a study involving two toasters. Seven users of varying levels of functional capability were observed using a simple toaster and a relatively complex toaster. The problems observed in this qualitative observational study are used to derive a deeper understanding of interaction problems that could arise, and also gain first-hand experience in working with disabled users. Further, an analytical evaluation of the simple toaster is carried out incorporating tools such as task analysis and state charts to represent the demands made on users. The results are used to better understand the issues faced by users with capability loss and the nature of results produced by analytical and empirical methods.

Chapter 7: Experiment: Four Consumer Products

This chapter reports on an experimental study designed to investigate the relationships between user capabilities, product demands and task performance in inclusive interaction. Nineteen users of various functional capability profiles performed tasks with a clock radio, mobile phone, blender and vacuum cleaner. Task performance measures such as task times, errors, and rated difficulty were obtained in addition to measures of sensory, cognitive and motor capability of each user. Capability-demand scatter plots and co-relational analysis were used to investigate the emergent relationships in the data.

Chapter 8: Discussion and Conclusions

Research results from the exploratory studies are drawn together in this final chapter and recommendations are made for further work in developing a data driven capability-demand interaction model as a basis for inclusive product evaluation.

Chapter 2 **Background**

2.1 Chapter Overview

In this chapter, a review of background literature relevant to the thesis research is presented. The chapter is organised around three key areas. Firstly, the concept of Inclusive Design and the nature of Inclusive Design problems are analysed. It is demonstrated that Inclusive Design requires knowledge of older and disabled user characteristics and that the process of Inclusive Design involves decision making, optimisation and trade-offs. The literature on the information needs of designers is also briefly reviewed to understand what types and formats of user data would be most useful and usable in actual design practice. Secondly, the literature and population data on disability are reviewed to understand the definitions and models underpinning the concept. After demonstrating the limitations of existing data, it is concluded that functional capacity or capability measures are required to support Inclusive Design.

Thirdly, a review of the literature on product evaluation methods is presented, drawing a distinction between empirical and analytical evaluation methods. State of the art inclusive analytical evaluation methods are reviewed with a discussion of their advantages and disadvantages. A case is made for further research into developing analytical evaluation methods for Inclusive Design. Finally, available literature and data on the characteristics of older and disabled users are considered with an eye toward supporting analytical product evaluation. It is concluded that there is currently a lack of coherent and comprehensive user capability data for product evaluation. In addition, the relationship between user capability measures and performance in real-world tasks needs further investigation and understanding. Thus, the chapter closes with the identification of a need for better understanding of how human capability data can support analytical product evaluation for Inclusive Design.

2.2 Inclusive Design

2.2.1 Inclusive Design Definition and Philosophy

Inclusive Design is defined as the: "Design of mainstream products and/or services that are accessible to, and usable by, people with the widest range of abilities within the widest range of situations without the need for special adaptation or design" (BSI, 2005). This design philosophy aims to consider the needs and capabilities of older and disabled people in the design process. It is focused on mainstream product design as opposed to assistive technology by avoiding aids, adaptations and stigmatising designs. Product aesthetics and desirability are also major concerns sitting alongside accessibility and usability attributes. Ideally Inclusive Design should not be viewed as a sub-speciality of design, but rather it should be perceived as being about good design (Coleman, 2006; Keates & Clarkson, 2003a).

Other terms such as 'Universal Design'(Story, 2001), 'Design-For-All' (Sims, 2003) and 'Trans-Generational Design' (Pirkl, 1994) have been used to encapsulate the approach of designing for variation in human capabilities and age generations. Universal Design is defined as "*the design of all products and environments to be usable by people of all ages and abilities, to the greatest extent possible*" (Story, 2001). As the term 'universal' may connote a 'one size fits all' approach, the term 'inclusive' declares the intent of the designer to maximise inclusion (Newell & Gregor, 2000). However, despite terminological differences, these approaches are fundamentally about the *accommodation of human diversity* in the design of products and services (Hosking et al., 2010; Story & Mueller, 2001). Seven principles of Universal Design were developed as a guide to carrying out Universal Design (Story, 2001) shown in Table 2-1.

Table 2-1. The Seven Principles of Universal Design

Principles of Universal Design
1. Equitable Use: The design is *useful* and *marketable* to people with *diverse abilities*
2. Flexibility in Use: The design accommodates a *wide range* of individual *preferences* and *abilities*.
3. Simple and intuitive: Use of the design is *easy to understand*, regardless of the user's experience, knowledge, language skills, or current concentration level.
4. Perceptible Information: The design communicates necessary information effectively to the user, regardless of ambient conditions or the user's sensory abilities.
5. Tolerance for Error: The design *minimizes hazards* and the adverse consequences of accidental or unintended actions.
6. Low Physical Effort: The design can be used *efficiently* and *comfortably* and with a minimum of fatigue.
7. Size and Space for Approach and Use: Appropriate size and space is provided for approach, reach, manipulation, and use regardless of user's body size, posture, or mobility.

These definitions of Inclusive Design and Universal Design suggest that products should firstly attempt to cater to the needs of users with reduced functional capacity. This implies that designers should be aware of and have information on the range of user functional capabilities with which to make design decisions and evaluate designs. Secondly, it involves a *maximisation* process implied by the phrase, ‘to the greatest extent possible’. Though a product cannot be inclusive or accessible in the absolute sense (Vanderheiden & Vanderheiden, 1992), it is useful to think about ways to make a given design more inclusive. These definitions represent an idealistic view that should be strived for, but also one that might be difficult to achieve in practice.

The Inclusive Design approach is a user centered approach to design (Keates & Clarkson, 2003a; Poulson, Ashby, & Richardson, 1996), where the fundamental premise is that accessible and usable products and services can only be developed by first *knowing* the intended users (Keates & Clarkson, 2003a; Mayhew, 1999; Nielsen, 1993). Knowledge of users refers to understanding exactly who will be using the product or service and what will be their capabilities, needs and preferences.

In a controversial article, Norman (2005a) argues that the user centred design approach may be harmful. He explained that a narrow view of users could result in designs that cater to specific user groups and lead to a fixation on designing for individual tasks or aspects of the interaction. Instead, he suggests a broader *activity centred* approach to design, whereby understanding user activities and their interrelationships are equally as important as understanding detailed user characteristics. Norman further argues that this extreme focus on individual users encapsulated in scenarios and personas might work well for individual screens and controls, but they could actually work against supporting a cohesive sequence of tasks that comprise a larger activity. Norman’s article spurred many debates which led to him publish a clarification on his website (Norman, 2005b). He further explained that User Centred Design had become limited to focusing on individual users at the expense of considering task flows and activity limitations. In addition, he advocated looking for common errors made in performing activities with an eye to designing them out or providing adequate assistance.

From an Inclusive Design standpoint, both the user centred and the activity centred perspectives are useful in understanding how to design products for a range of people to support their activities. Norman’s argument is essentially one of focus where he is attempting to warn designers not to focus on individual users alone, but rather consider the entire system in which users will act. Ergonomics theory describes the components of this system which comprises the user, product, environment and activities (Bridger, 2003; Karwowski, 2002). For effective Inclusive Design, each element in the system and their interactions warrant equal analysis. Designers are required to make design decisions that can support user activities while at the same time making sure that the maximum number of users could access and use the product.

Recognising that a completely inclusive product is an ideal as opposed to a practically achievable result, the focus of Inclusive Design should be on implementing a design process that gives due consideration to the aforementioned system. In essence, this embodies a *pro-active* approach to accommodating diversity in product design (Stary, 2000). The result of such a process should be improved product designs that minimise the exclusion of less capable populations.

2.2.2 Historical Context

The field of Ergonomics and Human Factors was born as a distinct profession after World War II ended in 1945 (Sanders & McCormick, 1993). There was a distinct need for the discipline given practical needs such as designing aircraft and ships to fit human users and general technological advances resulting in more complex engineered systems (Wickens & Hollands, 1999). Sub-fields of study including Anthropometrics (measurement of the human body) gave rise to ergonomic concepts such as designing for a *population of users* and the percentage of a population that could be accommodated by a given design (Bridger, 2003).

Measures of static anthropometry, functional anthropometry and forces exerted were collected over the years in databases and utilized by designers to accommodate 95% of the user population for a given dimension (with a Normal distribution of each measure). An example of such a data set is (Peebles & Norris, 1998). In general, the dimensions for 5th-percentile female to the 95th-percentile male were used as cut-off points for the extremes of the distribution (Pheasant, 1987). However, this approach was found to be problematic when multiple (multivariate) measures were required to assess the degree of fit. In actuality, there is no true 5^{th} or 95h percentile person (Pheasant, 1987). If a person is at the 95^{th} percentile for leg length, it does not mean that he or she is at the 95^{th} percentile for another measure such as stature. Therefore, if a set of 95^{th} percentile measures are used in calculating design accommodation, the result would be a design that accommodates far less than 95% of the user population (Pheasant, 1987). The need for integrated multivariate data sets soon became apparent (Porter et al., 2004).

The study of older and disabled users has always been a sub-specialty of the Ergonomics and Human Factors field. Compared to other areas of study, it historically remained a small research field resulting in limited data for heterogeneous populations (Kondraske, 2006d). However, Inclusive Design has recently gained prominence due to the unprecedented ageing of the world population (United Nations, 2009) and the expected increases in disability that is expected to accompany it (World Health Organisation, 2011). The responsibility is now on designers to develop products and services that could support a population where one person out of every five is aged 60 or over by the year 2050 (United Nations, 2009).

In addition to population trends, there has been a shift from a medical model of disability to a social model of disability where disability results not only from an impairment to body structures and functions, but also from the design of the built and manufactured environments (World Health Organisation, 2001). Years of lobbying Government have resulted in legislation that makes it mandatory for manufacturers and businesses to produce accessible products, environments and services. In the UK, the Disability Discrimination Act of 1995 (DDA) prohibited discrimination against disabled people in a range of areas including the provision of goods and facilities. The DDA has now been repealed and replaced by the Equality Act 2010. The new Act has introduced protection from three new forms of disability discrimination: (1) direct discrimination because of disability in relation to goods, facilities and services; (2) indirect disability discrimination, and (3) discrimination arising from disability (Government Equalities Office, 2010a; Government Equalities Office, 2010b). In the United States, there is the Americans with Disabilities Act (ADA) which has also been recently updated to make the Act more comprehensive and explicit with respect to serving customers with disabilities (U.S. Department of Justice, 2011).

Given these historical developments, Inclusive Design is poised to be the dominant design approach in the coming years to effectively satisfy the target user market while adhering to the legal guidelines for accessible environment and product design.

2.2.3 The Inclusive Design Process

Working definitions of Inclusive Design and Universal Design have been developed that acknowledge the constraints of a commercial environment on the design process. These pragmatic definitions of Inclusive Design are: "*An inclusively designed product should only exclude the users that the product requirements should exclude*" (Keates & Clarkson, 2003a) and "*The process of designing products so that they are usable by the widest range of people operating in the widest range of situations as is commercially practical*" (Vanderheiden, 2001). Vanderheiden (2001) argues that designing inclusively is about following a *design process* that takes place within the time and cost constraints of real world product development. Importantly, designers should be *aware* of the effects of design decisions on product accessibility and usability at every step in the design process.

In order to support manufacturers and designers in following an Inclusive Design process, the British Standards Institute (BSI) has published BS 7000-6:2005 as a standard document for implementing and managing an Inclusive Design process (BSI, 2005). Design methods and frameworks also exist to support designers including USERfit (Poulson et al., 1996) and the 7-level design approach to Inclusive Design (Keates & Clarkson, 2003a). These frameworks support the designer in addressing the practical and social acceptability of product designs for a range of users. Thus the literature provides some support for designers and manufacturers wishing to follow an Inclusive Design process.

2.2.4 Key Theoretical Concepts in Inclusive Design

2.2.4.1 The User Pyramid

Benktzon (1993) first provided the *user pyramid* approach to Inclusive Design for understanding the range of user capabilities (Figure 2-1).

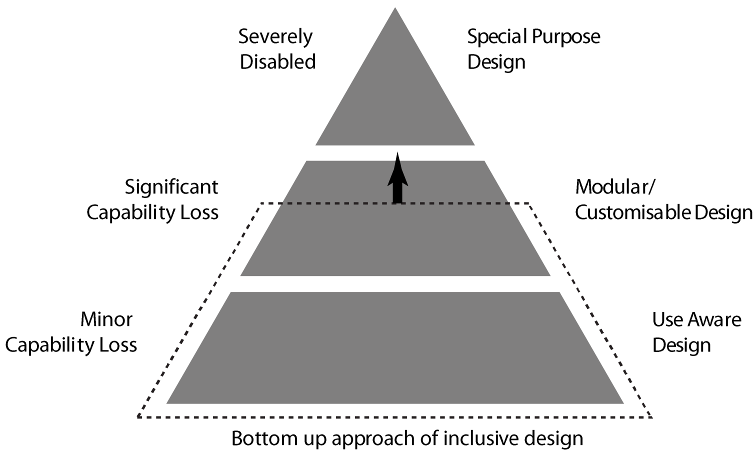


Figure 2-1 The user pyramid approach to Inclusive Design

The pyramid consists of three levels: (1) the lowest level of the pyramid represents people who are able bodied and with minor capability loss, (2) the second level represents people with significant capability loss such as people with mobility impairments and low vision, and (3) the third level represents people who are severely disabled and are unable to perform many of the activities of daily living without support. The Inclusive Design approach is basically 'bottom-up', where it strives to include users at higher levels in the pyramid. It also recognises that specialised designs and aids and adaptations would be necessary for users with moderate to severe capability loss (Etchell & Yelding, 2004).

2.2.4.2 The Inclusive Design Cube and the Concept of Design Exclusion

By definition, Inclusive Design requires an understanding of users at the group and population level (Carlsson et al., 2002; Keates & Clarkson, 2003a; Van der Vegte, 2002). In order to capture this population perspective, the user pyramid model was extended by Keates and Clarkson and the Inclusive Design Cube (IDC) was developed as shown in Figure 2-2 (Keates & Clarkson, 2003b). The cube model is a volumetric representation of the user population contained within three axes of sensory, cognitive and motor capabilities.

This model is based on the engineering model of the Model Human Processor (Card, Moran, & Newell, 1983), a representation of the capability range of individuals developed on a three-axis scale derived from the psychological dimensions of sensory, motor and cognitive capability. This provides a useful basis for an engineering model of human capability, even though the three dimensions are not independent and do interact in the performance of real world tasks. At one end of the cube, able bodied people are represented with high sensory, cognitive and motor capability. As one moves away from the able bodied point along the axes, the capability levels of the population decreases. The shaded volumes on the left cube represent proportions of the population that might benefit from different types of design solutions.

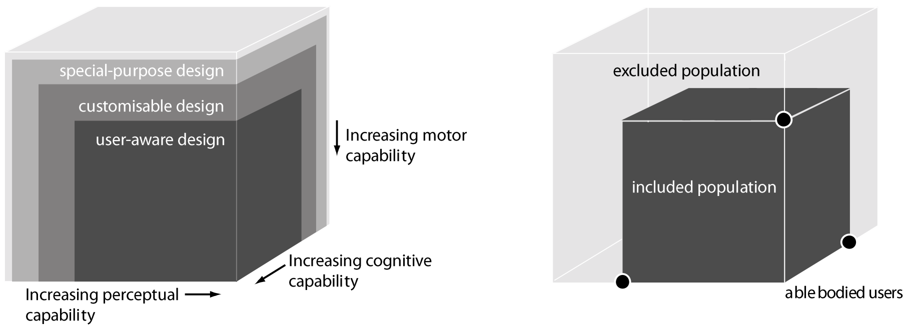


Figure 2-2 The Inclusive Design cube model proposed by Keates and Clarkson

Products can be viewed as placing demands on the sensory, cognitive and motor capabilities of the user population by setting demand levels on the three axes of the cube (Clarkson & Keates, 2003a; Clarkson & Keates, 2003b). The model is then useful in understanding the concept of design exclusion, where the user population contained within the dark cube on the right will be included and the population external to the dark cube will be excluded (Clarkson & Keates, 2003a; Clarkson & Keates, 2003b). Thus the aim of Inclusive Design is to minimise the volume of the theoretical excluded population while maximising the size of the theoretical included population.

2.2.4.3 Dynamic Diversity and User Sensitive Inclusive Design

Newell introduced the concepts of dynamic diversity and user sensitive Inclusive Design (Newell & Gregor, 2000; Newell & Gregor, 2002) that contributes to the theoretical base of Inclusive Design. Dynamic diversity acknowledges that the capability profiles of people in older and disabled populations are not only diverse, but they are also changing with time. He proposes the concept of user-sensitive Inclusive Design to reflect the conflicting nature of designing for different groups and also the difficulty in accommodating this dynamic diversity. In addition, one can consider ordinary users operating in *extra-ordinary environments* as being as disabled as *extra-ordinary users* operating in ordinary situations. The important point is the relationship between the user, the product and the operating environment ultimately leading to situations of disability.

Though the concept of dynamic diversity is an important conceptual tool for understanding the changing capability profiles of a population of users, it is more difficult to apply in practice because of a lack of time-varying user capability data. It might be possible to collect design relevant data via large scale longitudinal studies. However, challenges remain when considering people with certain conditions such as arthritis where physical capabilities can vary sometimes on a day to day basis.

2.2.4.4 Optimisation and Trade-Offs with Different User Groups

Because the definitions of Inclusive Design contain phrases such as “widest possible audience” and “all ages and abilities, to the greatest extent possible,” Inclusive Design can also be characterised as an optimisation process. In this case, the usability and accessibility of a given product needs to be maximised for the maximum number of people. Conversely, it would be desirable to minimise the number of people that the product design excludes. In optimising the product to support the maximum number of users, it invariably involves the navigation and resolution of trade-offs in accommodating different user groups (Goonetilleke et al., 2003; Gupta, Keates, & Clarkson, 2003).

2.2.5 Challenges and Limitations to Inclusive Design

As previously stated, absolute inclusivity is an ideal and it is unlikely that any given product can be inclusive in an absolute sense. The reality of the Inclusive Design approach is one of aiming to making products *more* inclusive and accessible (Vanderheiden & Vanderheiden, 1992). Problems can occur when accessibility for particular user groups is designed in a piecemeal manner, resulting in other user groups being partially satisfied. Changes made to the product to accommodate one type of disability might disadvantage people with another type of disability (Vanderheiden & Vanderheiden, 1992). Vanderheiden suggested that a *holistic* approach is required where the entire product must be *completely* accessible for certain user groups rather than the product partially satisfying different user groups.

This problem was also highlighted by Newell in outlining the approach of user sensitive Inclusive Design (Newell & Gregor, 2000; Newell & Gregor, 2002). Newell pointed out that designing a product for people with particular types of disability can make the product more difficult to use by people without disabilities and also for people with different types of disability. He suggests that the 'excellent is the enemy of the good' and that 'accessibility by all' could provide a barrier to improved 'accessibility by most.' This sentiment is echoed by other researchers in the field (Hawthorn, 2003), leading to the conclusion that Inclusive Design is not a simple problem that is easily addressed, but one that requires thoughtful consideration of the impact of the range of human capability loss on design.

Another problem with Inclusive Design occurs where there is a tension between an attempt to focus on and concretise the target users in the designer's mind while at the same time trying to consider designing for a heterogeneous population of users. In the field of marketing, the technique of market segmentation is used to clearly define specific groups of people by lifestyle variables. This provides the designer with specific information on how to design for a specific target market (Bellerby & Davis, 2003). Designers use techniques such as personas in narrowing down the vague concept of the 'user' or 'users' into concrete characters that can be used to aid design (Cooper, 1999; Grudin & Pruitt, 2002). Inclusive Design as an all encompassing approach tends to move in the opposite direction and expand the design space to a wide range of users of different ages and capabilities. Therefore, for practical implementation, Inclusive Design has to work within business constraints and minimise user exclusion within the target market segments (Hosking et al., 2010; Hyppönen, 1999; Keates & Clarkson, 2003a).

Market segmentation seeks to sub-categorise a target market into several groups with the sub-segments described by different preferences for products and services and methods of delivery (Moschis, 1992). It is based on the assumption that individuals can differ on various dimensions including perceptions, attitudes and consumption behaviour (Moschis, 1992). The aim of segmentation is to find an optimal set of segments that represent similarities within the segment, but also differences between segments.

One popular segmentation model by Moschis is his Gerontographic model consisting of four segments: (1) Healthy Indulgers, (2) Healthy Hermits, (3) Ailing Outgoers and (4) Frail Recluses (Moschis, 1992). Healthy Indulgers are in good health, independent and socially engaged. Healthy Hermits are in good health, but they are psychologically and socially withdrawn. Ailing Outgoers tend to have poor health but are socially active. Frail Recluses are in poor health and also socially isolated. Older people may move from one segment to another depending on the trajectory of ageing (Moschis, 2003). Other segmentation models of the older population are available (Snyder, 2002) in the marketing literature, and more are projected to become available as the older market expands.

Though challenging, the design of a few inclusive products provides encouragement (Coleman, 2006). Products such as the OXO Good Grips line of kitchen and garden tools, the BT Big Button and Freestyle phones, and the Ford Focus demonstrate mainstream Inclusive Design success (Cardoso, 2005; Coleman, 2006; Hosking et al., 2010; Warburton, 2005). Designers can embrace constraints and challenges, resulting in greater levels of creativity and innovative product designs (Cassim, 2004; Etchell & Yelding, 2004).

2.2.6 Supporting Designers

Designers require adequate user information if successful Inclusive Design is to be achieved (Clarkson & Keates, 2003b; Goodman-Deane et al., 2010). In providing designers with end user information, there are two main considerations. On the one hand, valid data on user needs, capabilities and preferences are required. On the other hand, this information needs to be filtered and presented for *use* by designers in the design process (Wilcox, 2007). Thus both of these requirements need to be addressed if designers are to utilise user information. An information quality perspective could be useful where the user data is evaluated in terms of the intrinsic, contextual, representational, and accessibility data quality (Persad & Clarkson, 2005).

Several industry surveys have been carried out in order to determine the needs of designers and manufacturers. Studies by Vanderheiden (Vanderheiden & Tobias, 2000), Sims (Sims, 2003), Dong (Dong, 2004) and Goodman-Deane (Goodman-Deane et al., 2010; Goodman, Dong, Langdon, & Clarkson, 2006a; Goodman, Langdon, & Clarkson, 2007; Goodman, Langdon, & Clarkson, 2006b) demonstrate recurring themes in terms of the barriers and requirements for the implementation of Inclusive Design in industry. Major barriers include the lack of time and budget, lack of knowledge and tools and lack of a justifiable business case (Dong, 2004; Goodman-Deane et al., 2010; Goodman et al., 2006a). In order to overcome these barriers, the following guidelines are suggested:

Support the designer workflow with user information: In order to design inclusively, knowledge on the spectrum of human sensory, cognitive and motor capabilities is required. The fields of Biology, Psychology, Medicine, Occupational Therapy and Ergonomics all contribute data and information to the Inclusive Design knowledge base. In addition, data on the proportions of people with various capability profiles is required (Carlsson et al., 2002), with information on the implications of these capability profiles on product interaction (Johnson et al., 2010). In supporting designers with human sensory, cognitive and motor capability data, consideration must be given to aiding the designer in learning about unfamiliar aspects of ageing and disability (Bellerby & Davis, 2003). This is necessary so that designers could access and utilise user information in their design activities.

Provide just-in-time information: It has been suggested in the literature that “just in time” information is needed to support designers when working on Inclusive Design problems (Keates & Clarkson, 2003c). In a time pressed design environment, quick and dirty or ‘lightweight’ methods will be favoured as opposed to methods that require significant time and effort (Goodman et al., 2006b). To provide just in time information, designers would like it to be contextual and relevant to the product they are designing at the moment. The design of knowledge bases to support Inclusive Design should not only consider supporting the designer workflow, but also support ease of learning and finding new information quickly and effectively. The ability to find up-to-date and usable data on user capabilities is also necessary. Tools and methods that support the Inclusive Design process might not be used if they require large investments of time and are complicated to learn (Wilcox, 2007).

Providing inspiration and avoiding prescription: Traditionally, human factors information is delivered in the form of guidelines, handbooks and tables of data (Pirkl & Babic, 1988a; Poulson et al., 1996; Smith et al., 2000; Vanderheiden & Vanderheiden, 1992). However, text heavy design information and tables of numbers are difficult for designers to use (Wilcox, 2007). In addition, information that is written in an academic style and information that is out of date will be disliked (Goodman et al., 2007). Designers prefer information that is accessible, concise, visual, up to date and easy to use (Goodman et al., 2007; Wilcox, 2007). Designers prefer design information and guidance that is not overly prescriptive (Goodman et al., 2007). Instead, they prefer information that could provide design inspiration and also allow for the freedom to make their own design decisions (Dong, 2004; Goodman et al., 2007; Wilcox, 2007).

Both qualitative and quantitative user data are required by designers: Qualitative data such as user profiles, stories, videos and other multimedia can be useful for generating insights and developing empathy for users (Goodman et al., 2007). Quantitative data is also required on numbers of people with various profiles of functional capability loss in order to make estimates of design exclusion (Goodman et al., 2007). Quantitative data is also useful to business managers and decision makers for understanding the sizes of market segments and forecasting increased revenues by accommodating users with reduced functional capacity (Brinck, 2005; Dong, 2004; Hosking et al., 2010; Waller, Langdon, & Clarkson, 2010b).

In a recent triangulated study into information use by designers, Goodman-Deane (2010) described a framework of four areas for consideration: the influence of the client (and the design brief), the informality of methods and tools, the variation in design processes and methods used, and the effect of time and cost constraints. The key implication of these findings is that there needs to be a body of methods, tools and data for Inclusive Design upon which designers and clients could draw as needed. The many constraints of a real-world environment result in quick and informal methods being the preferred choice. Therefore, special attention needs to be paid to the *usability* of methods and tools for successful adoption (Wilcox, 2007). There is also scope for the further development of Inclusive Design materials that could assist the designer through all phases of the design process. One such resource that aims to bring the different types of Inclusive Design information together is the Inclusive Design Toolkit (Cambridge Engineering Design Centre, 2011; Clarkson, Coleman, Hosking, & Waller, 2007).

Keates and Clarkson developed the Inclusive Design knowledge loop (Figure 2-3) as a framework for representing the iterative activities and information flow in the design process (Keates & Clarkson, 2003a). The loops shows the capture of user data and information in various representations which serve to inform designers and business managers about the target user population. This information can be used throughout the design process by relevant stakeholders (contained in the ‘information users’ circle on the diagram). Clients can reference user data in their briefs while designers can use user data in design activities. Once the user data has been used to design inclusive products and services, the products are validated by end users to ensure that they meet users’ needs.

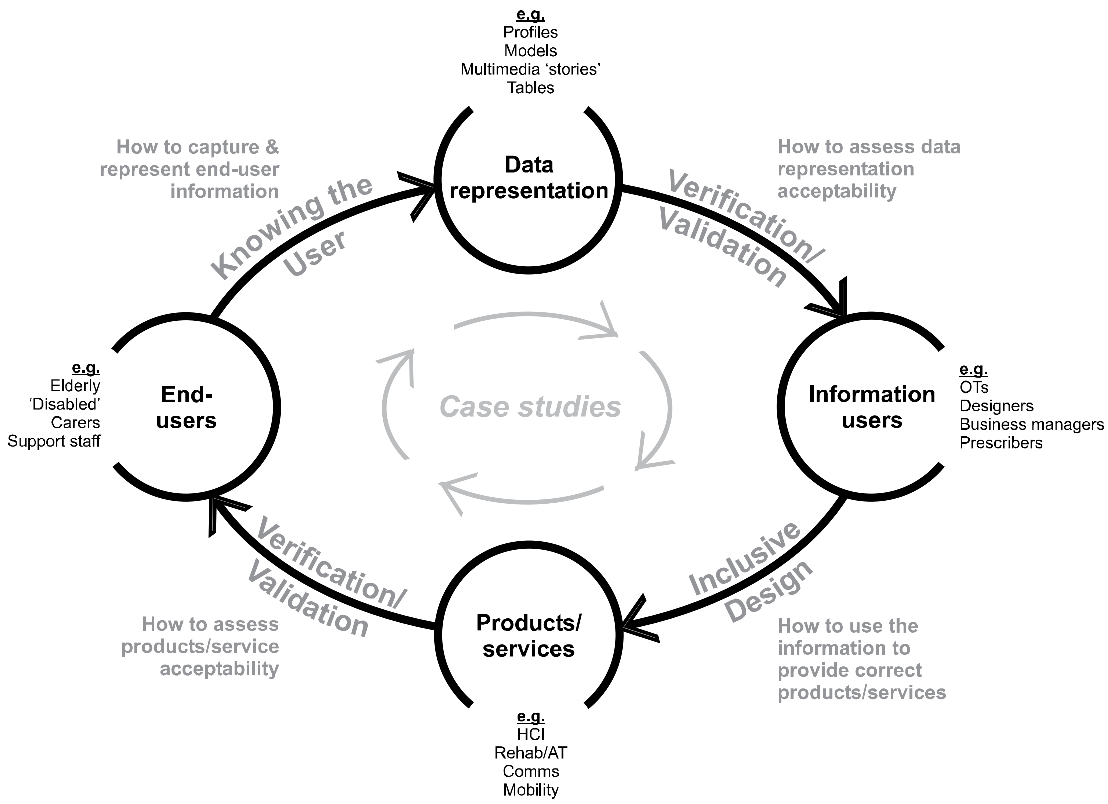


Figure 2-3 The Knowledge Loop presented by Keates and Clarkson (2003)

2.2.6.1 Supporting Decision Making

Product development can be viewed within a framework of decision-making (Dorst, 2003; Krishnan & Ulrich, 2001; Wilcox, 2007). The decision perspective is particularly relevant to Inclusive Design because all decisions, conscious or unconscious, impact on the quality of the final design. Decisions about the functionality, features and industrial design of consumer products must be made considering the impact on end users. Thus adequate user information is required to make informed design decisions as an integral part of the Inclusive Design process (Stary, 2000). Benyon et al. argue for a process of 'inclusivity analysis' to become standard product design practice together with the understanding and justification of design decisions on excluded users (Benyon, Crerar, & Wilkinson, 2001). Designers are required to make different types of decisions when designing a given product. For example, consider the two blender designs in Figure 2-4.

Figure 2-4 Two blender designs with different speed controls

The first type of decision involves deciding on the particular type of interface feature required to expose product functionality. Two options for controlling the blender motor speed are apparent - a rotary control on the left and a series of push buttons on the right. The rotary control demands a grasping and rotational action while the push button control demands a finger push action. Since both types of control are suitable for the operation at hand (changing the motor speed), the designer must evaluate the two types of controls in terms of ease of operation – especially for users with reduced hand function.

Though human factors literature contains guidance on the type of interface feature that would be suitable for a particular control function (Sanders & McCormick, 1993; Woodson, Tillman, & Tillman, 1992), in most cases the guidance does not extend to users with functional capability loss. In addition, there may be conflicting recommendations given in guidelines for accommodating different forms of functional loss and accommodating one type of disability might cause problems for users with another type of disability (Vanderheiden & Vanderheiden, 1992). For example, replacing visual output with auditory output for users with visual problems might work for users with visual impairments, but it would cause problems for users with hearing impairments. Another example would be changing slide controls to make them easier to activate while simultaneously increasing the probability of accidental activation (Electronic Industries Alliance & Electronic Industries Foundation, 1996). In light of this, it is near impossible to follow all design guidelines and recommendations simultaneously (Vanderheiden & Vanderheiden, 1992).

Table 2-2 Two types of decisions in designing the product interface

Decision Type	Demand on user	Example
Choosing among alternatives for interface features	Action demand e.g. Push action or rotate action?	Choice between a push button or a rotary control
Setting the specific attributes of interface features	Performance demand e.g. size, force, speed	Push button force, diameter Rotary control torque, diameter, thickness

Once the type of interface feature has been selected, the second type of decision involves the setting of specific interface attributes such as dimensions and forces required for operation, i.e. detail design. These decisions result in specific performance demands on the user such as the fit of the control to the finger/hand and the level of force the user must exert in order to activate the control. The designer requires user information that can aid in making both types of decisions (Table 2-2) in the Inclusive Design process.

2.2.6.2 Supporting the Resolution of Trade-Offs

Inclusive Design problems can present unique challenges when generating strategies to accommodate the sometimes conflicting needs of diverse groups of users. For example, designing to accommodate one user group could affect the usability of the product for another. Hawthorn describes the practicalities of designing a usable email client for older users that resulted in usability problems for general and expert users (Hawthorn, 2003). He suggests that the effects of ageing on physical, sensory and cognitive capabilities need to be considered in combination when designing for older users. He concludes that the interface techniques used for creating a successful interface for older people are "*directly at variance with the techniques used by designers for supporting a modern, feature rich application*."

As another example, making design changes to accommodate users with vision loss by providing audio output might cause problems for users with hearing loss. Providing feedback in both visual and auditory modalities may not satisfy users with both vision and hearing problems. It is argued in the literature that good design for older people will result in good design for all users (Fisk, Rogers, Charness, Czaja, & Sharit, 2004, p. 147). These contradictory examples challenge that assumption and show that this may not always be the case due to the interaction of accessibility, usability and satisfaction. Younger and expert users may prefer a more feature rich product and they could become frustrated with overly simplified designs.

The literature contains few examples of support systems for assisting the designer in visualising and making trade-offs based on user characteristics. Zajieck describes a methodology for designing speech systems for older adults using a pattern language (Zajicek, 2004). Each design pattern includes an explicit trade-off section that describes how the proposed solution to a given interface problem could cause other usability problems in the interface. Methods such as pattern languages might provide designers with support for understanding and handling trade-offs in the Inclusive Design process.

2.2.7 Section Summary

In summary, Inclusive Design requires knowledge about the spectrum of human capability across a target population. Designing inclusively can be understood as a decision-making, optimisation and trade-off process. In addition, designers and manufacturers require design methods, tools and data (both qualitative and quantitative) that could be easily applied at different stages in the design process.

2.3 Understanding Users with Disabilities

As previously mentioned, Inclusive Design rejects the notion of 'the average user' as a myth and embraces the diversity in user characteristics. Therefore, rather than adopt standard user-centred design approaches investigating average capabilities and population norms, the Inclusive Design approach seeks to map the product to a space of capabilities ranging from low to high, and to systematically minimise product demands to achieve a more accessible product. Understanding this variation in sensory, cognitive and motor ability is therefore an essential first step in designing effective product evaluation systems. Based on a list originally presented by Norris (Norris & Wilson, 1997), Table 2-3 shows a categorisation of relevant user characteristics that are necessary for consideration when designing for people. These factors must all be considered when designing consumer products intended for general use.

In addition to sensory, cognitive and motor capabilities, personal and socio-economic factors also need to be taken into account. User motivation and values play an important role in the use of products and services, and these affect the way that the product is perceived (Crilly, Moultrie, & Clarkson, 2004; Crilly, Moultrie, & Clarkson, 2009). Aesthetics and usability are inter-related as users perceive attractive products to work better (Norman, 2005a). Depending on the target market, users will respond differently to advertising and make different purchasing decisions (Moschis, 2003). Since this is the first point of access of a product, personal and socio-economic factors are critically important. These factors are also important considerations in the context of use of various products. They determine user behaviour when interacting with the product and the type of coping strategies that are employed (Yoxall, Langley, Musslewhite, Rodriguez-Falcon, & Rowson, 2010c).

Importantly, the factors listed in the table are not static, but rather they change with time throughout the life course. The *Health Status/Disability* dimension can impact the five other dimensions in temporary or permanent ways. This gives rise to a complex picture of the user which consists of many interacting factors in a real world context. To fully understand the user would mean to fully understand the target population along all of the dimensions. Therefore this categorisation highlights the complexity and wide range of consideration that must be given when designing inclusively.

Table 2-3 User Characteristics Relevant to Product Design

Domain	Considerations
1. Sensory Capabilities	• Vision, hearing, taste, smell, touch, balance, kinaesthesia • Aids for vision and hearing
2. Cognitive Capabilities	• Information processing capacity and performance (perception, working memory, attention, long term memory) • High level cognition: reasoning, planning, problem solving • Knowledge (declarative and procedural), experience and mental models of products
3. Motor Capabilities	• Static (body dimensions) and functional anthropometry (range of movement) • Strength with various body postures • Fine motor skills (hand-eye coordination, dexterity) • Gross motor skills (body movement, locomotion) • Endurance and stamina • Aids for movement
4. Personal and cultural factors	• Preferences • Aspirations • Motivation • Language
5. Socio-economic factors	• Age • Gender • Income • Lifestyle • Education • Family support
6. Health Status/Disability	• Medical Conditions (short and long term effects) • Accident/injury (short and long term effects) • Medication effects • Pregnancy

2.3.1 Understanding Disability: The ICF Model

The concept of disability needs to be understood as a precursor to Inclusive Design, as many people with functional capability loss are said to have one or more *'disabilities'*. In the field of disability studies, there are complications in rigidly defining disability because *"disability is a complicated, multidimensional concept"* (Altman, 2001). The World Health Organisation (World Health Organisation, 2001) defines disability as *"Any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being."* For further information on the development of disability definitions, classification schemes, and models up to the current version of the International Classification of Functioning, Disability and Health, the reader is referred to (Altman, 2001).

The International Classification of Disability, Functioning and Health (ICF) (World Health Organisation, 2001) and the International Statistical Classification of Diseases and Related Health Problems (ICD-10) (World Health Organisation, 1992) together provide a comprehensive language for disability and health problems. The ICF shows a move from a *medical model* of disability toward a *social model* of disability. The medical model suggests that disability is a trait of the individual, while the social model suggests that disability arises because of the mismatch between the individual and the social and physical environment. This means that disability is caused by the demands of the environment rather than being a specific attribute of the person. Thus a person is disabled only when viewed in the context of the environment in which he or she must function.

The model of disability given by the World Health Organisation (WHO) is shown in Figure 2-5 (World Health Organisation, 2001). This model highlights the complexity and interaction between the concepts of medical and social models of disability. Various health conditions impact on a person's body structures and functions, thus affecting the person's activities and participation in the social environment. Environmental and personal factors also contribute to the person's body functions, activities and participation, resulting in a complex model of inter-relating factors. The ICF model is now widely adopted and used as the basis for the collection of data on disability for medical and decision-making purposes. Since it recognises the complexity of disability as a transactional process between multiple factors, it is an appropriate model for understanding and framing disability issues.

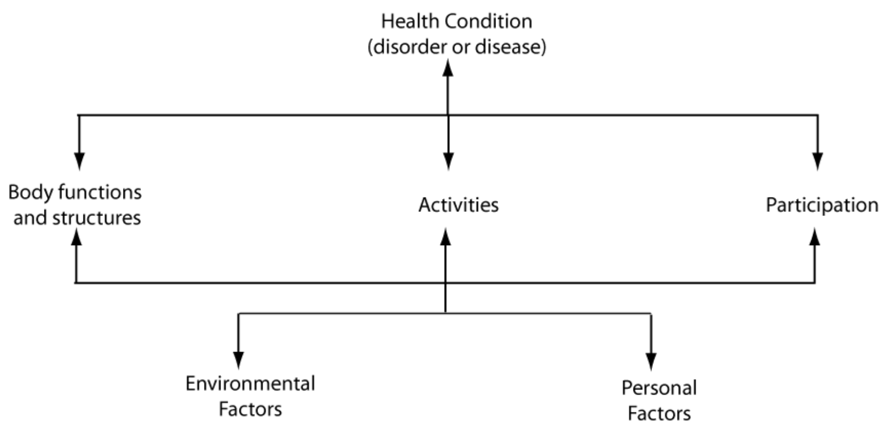


Figure 2-5 The model proposed by the World Health Organisation for Disability in the International Classification of disability, functioning and health

From the ICF model, definitions for common terms are set out as shown in Table 2-4.

Table 2-4 Definitions of Disability and related concepts (World Health Organisation, 2001)

Concept	Definition
Disability	Any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being.
Body Functions	Body Functions are physiological functions of body systems (including psychological functions).
Body Structures	Body Structures are anatomical parts of the body such as organs, limbs and their components.
Impairments	Impairments are problems in body function or structure such as a significant deviation or loss.
Activity	Activity is the execution of a task or action by an individual.
Participation	Participation is involvement in a life situation.
Activity Limitations	Activity Limitations are difficulties an individual may have in executing activities.
Participation Restrictions	Participation Restrictions are problems an individual may experience in involvement in life situations.
Environmental Factors	Environmental Factors make up the physical, social and attitudinal environment in which people live and conduct their lives.

From these definitions, though a user may have impairments, disability occurs when a user cannot perform a required activity. Thus from a design perspective, if the product and environment is changed to accommodate the user's capability, the apparent disability would be removed. This analysis leads to the term *'disabling by design'*, as it is within the designer's power to create products and environments that accommodate people who may not have 'average' capability.

From an Inclusive Design perspective, the ICF model of disability demonstrates that disability is a continuum rather than a binary category. To consider a person as being disabled or not-disabled is a gross simplification of disability, ignoring the many levels and combinations of sensory, cognitive and motor functional loss that can occur. The ICF model also forms the basis of an extensive taxonomy of body functions, structures, activities and participation that is very comprehensive and medically oriented. A few of these aspects are relevant to product design, however in the most part, such classifications are unsuitable for general design purposes (Carlsson et al., 2002).

2.3.2 Disability Data

In reviewing disability definitions and statistics on the prevalence of disability in the UK population, a major issue was discovered. There are no "Gold Standard" definitions of disability due to the fact that disability is a multidimensional and dynamic concept (Bajekal, Harries, Breman, & Woodfield, 2004; Tibble, 2004). Because of this, the population estimates of disability from different surveys are not directly comparable, and users of these disability estimates are required to choose a survey estimate based on their particular needs (Bajekal et al., 2004). This leads to problems when trying to collate and compare disability prevalence data from different sources.

Fujiura presents an overview of Disability Measurement Systems in the Handbook of Disability (Fujiura & Rutkowski-Kmitta, 2001). Essentially, there are three sources of population-level disability data: national censuses, household surveys and administrative registries. Because samples for censuses and household surveys are aimed at statistical accuracy, the data collection is costly in terms of labour and resources. This in turn limits the depth and detail of disability measurement instruments, which means that the resulting data is limited to a few indicative measures that are relatively easy to collect. For the UK, Bajekal presents an overview of disability surveys available (Bajekal et al., 2004) and the context of their estimations based on the definition of disability used. The UK Department of Work and Pensions also gives a table of disability prevalence from different surveys (Tibble, 2004) which stands at 22% of the UK population using the Disability Discrimination Act (DDA) as the definition of Disability.

From a public health and medical standpoint, surveys with disability elements tend to measure the performance of individuals on tasks such as Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) with the aims of improving health and quality of life (Bajekal et al., 2004). Though ADLs and IADLs include activities such as eating, bathing, shopping and manageing money, measurements with respect to specific product interface features are usually not available. As disability surveys provide data for policy-making purposes or for population health purposes, the field of Human Factors and Ergonomics should aim to characterise and measure the capability of disabled users on a population level, specifically to support Inclusive Design.

Another problem with population data is that it quickly becomes outdates due to changing demographics (Yoxall et al., 2006). For example, 95^{th} percentile strength data for older adults collected in the past would not represent current 95^{th} percentile strength data in the current context of rapid population change. The only way this can be remedied is via standardised measurement at regular intervals together with statistical estimation tools that would enable better estimates of true population data.

2.3.2.1 Focusing on Functional Capacity for Design

Disability can be characterised by disease (etiology), onset, progressive nature (time varying) and severity (Sesto, Vanderheiden, & Radwin, 2004). Even though ageing, disease and trauma account for limitations in functional capacity, information on the causes of functional ability loss are less relevant than the actual levels of functional ability. It is more important to understand what a user can and cannot do, and what levels of difficulty are experienced, rather than focusing on the cause of the disability (medical model) (Cardoso, 2005; Sesto et al., 2004). However, at times it may be useful to understand the causes of disability such as medical conditions, trauma and medication effects in order to explore contextual and lifestyle factors of a particular user group (Persad et al., 2005).

Sesto also puts forward the concept of 'functional equivalence' where different medical conditions can cause similar problems in functional limitations, and common solutions could be found to satisfy these groups of conditions. Thus capability (or functional capacity) characterisations appear to be the most appropriate for product design and evaluation because information on what users can and cannot do could be directly translated into improvements in interface features (Carlsson et al., 2002; Jacko & Vitense, 2001; Sesto et al., 2004).

2.3.3 Section Summary

In this section, the definition of disability and the ICF model were reviewed. It was shown that population representative disability surveys are limited in the information that they could provide, and the concept of capability or functional capacity should be used to capture data on disabled populations. In the next section, product evaluation methods for Inclusive Design are reviewed.

2.4 Product Evaluation for Inclusive Design

In this section, basic product evaluation concepts and evaluation methods are reviewed. The theories of compatibility and capability-demand relationships are shown to form the conceptual basis for evaluating consumer products. Evaluation methods are categorised into empirical and analytical methods, and current analytical methods for Inclusive Design are analysed. An argument is made for the further development of analytical methods for Inclusive Design.

2.4.1 Product Evaluation Theory

Human Factors and Ergonomics theory describes four main system components when considering the interaction of people with designed products. These are: (1) the user, (2) the product, (3) the environmental context and (4) the activities and tasks over time that constitute the interaction (Bridger, 2003; Karwowski, 2002). The aim is to evaluate the degree of fit between users and the designed product by utilising various measures of compatibility. This assessment of compatibility can be conducted at various levels including the sensory, motor and cognitive levels of human functioning (Bridger, 2003; Karwowski, 2002). In addition, the concepts of *user capability* and *product demand* provide a useful framework for analysing user-product compatibility (Bridger, 2003; Clarkson & Keates, 2003b). This entails a comparison of the sensory, cognitive and motor demands of a product in relation to the capability levels of the expected user population (Clarkson & Keates, 2003b).

Wickens and Hollands define four types of measures that can be used to assess the level of compatibility between user and product (Wickens & Hollands, 1999). These measures are shown in Table 2-5 and include measures of speed/time, measures of accuracy/error, measures of workload and measures of preference. Measures of speed and accuracy can be observed and recorded objectively by an evaluator. Measures of workload and preference mostly rely upon subjective self reports. They also define evaluative measures or figures of merit as a derived measure from these four types of raw performance data. This can give “good” and “bad” endpoints that might be of great use to the designer of a product (Wickens & Hollands, 1999). For Inclusive Design, the focus is on deriving evaluative measures of the product that can indicate how well the design accommodates people with lower than average capability in sensory, cognitive and motor functioning.

Table 2-5 Measures of performance in human factors after Wickens (1999)

Measure Type	Question	Measure Type
Measures of speed or time	How long does it take?	Objective – Observe and record time taken
Measures of accuracy or error	How many errors are made?	Objective – Observe and classify errors made
Measures of workload or capacity demands	How difficult is it?	Subjective or objective – Self-Report or physiological measures
Measures of preference	What do you like or prefer?	Subjective – Self Report

Various evaluation methods are employed in order to measure the level of compatibility between the user and the product, and to derive the above evaluative measures. The field of human factors/ergonomics contains a range of such methods (Leonard, Jacko, Yi, & Sainfort, 2006; Poulson et al., 1996; Stanton, Hedge, Brookhuis, Salas, & Hendrick, 2004; Stanton, Salmon, walker, Baber, & Jenkins, 2005). These methods can be classified along various dimensions including time, cost, resources required, level of human factors expertise required, stages in the design process that the method might be the most useful, and the resulting output or prediction of the method (quantitative or qualitative).

Quantitative methods comprise the collection of data in a numerical format which is amenable to statistical analysis and hypothesis testing (positivist based) (Frankfort-Nachmias & Nachmias, 2000; Robson, 2002). Quantitative research usually allows the researcher to make inferences and generalisations about a larger population of interest based on a (random) sample from the population. Examples include surveys with various forms of quantitative measurement (Leonard et al., 2006). Qualitative methods result in data that are generally non-numerical in nature, for example narratives, stories and pictures. Sample sizes tend to be smaller that quantitative methods, and rather than aim for statistical generalisability, qualitative methods aim to capture the richness of the issues being investigated. Methods that fall in this category include interviews, observations and ethnographic research (Frankfort-Nachmias & Nachmias, 2000; Robson, 2002).

Quantitative and qualitative methods are not mutually exclusive, as some forms of qualitative data could be converted into quantitative data. For example qualitative data captured as user problems in video observations could be converted into counts and categories of problems (or errors) that are amenable to statistical analysis. It has been argued that there is a place for both quantitative and qualitative methods in the researcher's toolbox in hybrid strategies (Hignett & Wilson, 2004; Kanis, 2003; Robson, 2002). Qualitative methods can address shortcoming in quantitative methods and vice versa. For example, quantitative methods are likely to miss the richness and subtleties of individual human behaviour. Kanis argues that "rich descriptions of distinguished cases, rather than findings ironed out across individual participants to blurred and contextually stripped averages and corresponding standard deviations" would be more beneficial to designers in creating novel design solutions (Kanis, 2003). Qualitative observations of small samples can result in a deep understanding of important issues before larger quantitative studies are carried out (Cardoso, 2005). For the purposes of this review, an empirical versus analytical classification will be further examined.

2.4.2 Empirical Versus Analytical Evaluation Methods

Hartson et. al. describe a useful categorisation of product evaluation methods into analytical methods and empirical methods (Hartson, Andre, & Williges, 2001). Empirical evaluation methods measure design performance in actual usage scenarios by having users perform tasks with a product. Various performance metrics can be recorded such as time taken, number and type of errors, and subjective impressions. Thus the key feature of empirical methods is that they *directly involve the user*. Examples of such methods include user observation and user testing (Cardoso, 2003; Rosson & Carroll, 2002; Stanton et al., 2004; Stanton et al., 2005). On the other hand, analytical methods are based on the inspection and analysis of product features by utilising expert judgement, heuristics, empathic simulation, user capability data and predictive engineering models of user behaviour (Cardoso, 2003; Hartson et al., 2001; John & Kieras, 1994; Rosson & Carroll, 2002). Therefore, analytical methods rely on the analysis of the product under scenarios of use *without direct user involvement*. A balance of both analytical and empirical evaluation methods should be considered in the design process, depending on the resources available.

Analytical methods can be advantageous in that they require fewer resources than empirical methods including less logistical difficulties, reduced time and reduced cost (Gyi et al., 2004). Such methods can produce quantitative predictions of design accommodation or exclusion, and can support the business case given the quantitative nature of the results (Clarkson et al., 2003b; Gyi et al., 2004). They can also be conducted at various levels of detail as needed in a 'just-in-time' fashion (Keates & Clarkson, 2003c). However, there are also various disadvantages to analytical methods compared to empirical methods. Since designers do not come in direct contact with users, the richness of real-world interaction is lost and an opportunity is missed for generating empathy for users. Analytical methods also require 'standard' assumptions about how people interact, and opportunities for design might be missed. For example, various coping strategies employed by disabled users might trigger new design solutions. Analytical methods are based on interaction theories valid for homogenous user groups, which may prove to be invalid for heterogeneous user groups. Finally, analytical methods can lead to an over reliance on numbers for making design decisions, and if the resulting numbers are invalid or misapplied, there could be serious design consequences.

Though the importance of user involvement via empirical methods cannot be overemphasised, there is a place for analytical methods in the evaluation process due to the aforementioned constraints of time, cost and logistical difficulties in recruiting and testing with real users (Gyi et al., 2004). However, it is argued that analytical methods are particularly advantageous for Inclusive Design when the population of users that are accommodated or excluded has to be determined (Carlsson et al., 2002; Persad et al., 2005). Such a population approach is necessary because traditional sampling strategies for empirical studies cannot account for the various types and combinations of sensory, cognitive and motor capability loss (Carlsson et al., 2002; Feeney, Chrisholm, Petherick, & Summerskill, 1998).

This means that traditional user testing cannot hope to capture all the types of users with all types of problems for a given product design. The alternative is therefore to utilise analytical methods that are based on *capability data* of the target user population, and to estimate the inclusion/exclusion of various subpopulations based on this data. Through this method, the goal of broad based *inclusivity analysis* will be achieved. Estimations of included/excluded populations are only practically possible when sufficient user capability data exists with corresponding methods for applying it. Therefore, research effort is needed for supporting analytical inclusive evaluation of consumer products.

2.4.3 Examining Analytical Evaluation Methods

Table 2-6 lists a five part categorisation of various analytical evaluation methods. These are (1) guidelines, (2) empathic simulation, (3) expert evaluation, (4) predictive evaluation with user models and (5) predictive evaluation with databases of capability data.

Table 2-6 Analytical Evaluation Methods

Method	Examples
1. Heuristics, guidelines and checklists	Published guidelines, checklists and heuristics (rules of thumb) are used to evaluate a design and identify areas for improvement. E.g. (Carmichael, 1999; Fisk et al., 2004; Pirkl & Babic, 1988a; Pirkl & Babic, 1988b; Poulson et al., 1996; Story, Muller, & Mace, 1998; Vanderheiden & Vanderheiden, 1992)
2. Empathic Simulation	By utilising various simulation strategies, designers can experience what it is like to have sensory and motor capability losses. E.g. the Ford 3rd Age Suit (Hitchcock, Lockyer, Cook, & Quigley, 2001), Age Explorer (Evamy & Roberts, 2004), EDC Simulation Suit (Cardoso & Clarkson, 2006) and Simulation Software (Goodman-Deane, Langdon, Clarkson, Caldwell, & Sarhan, 2007).
3. Expert evaluation, Expert judgement	Expert judgement is used to evaluate a design and provide recommendations for improvement. Usually more than one expert is used to maximise the capture of potential problems (Cardoso, 2003).
4. Predictive Evaluation: User models	User models are used to predict various performance parameters of the interaction. E.g. GOMS, EPIC, ACT-R, and SOAR (Kieras, 2003).
5. Predictive Evaluation: Virtual user trials with capability databases	Human capability data is captured in a database and used to predict performance parameters and populations accommodated by a given design. Inputs include product design parameters and task parameters. E.g. HADRIAN (Porter et al., 2004), Exclusion Audit (Cardoso, 2005), Anthropometric Databases (Smith et al., 2000).

Heuristics, guidelines and checklists are quick and easy methods that can be applied by the designer to evaluate a product design (Carmichael, 1999; Fisk et al., 2004; Pirkl & Babic, 1988a; Pirkl & Babic, 1988b; Poulson et al., 1996; Story et al., 1998; Vanderheiden & Vanderheiden, 1992). These methods are not resource intensive and this is the dominant format available to designers. However, guidelines can be quite text heavy and too many could overload the designer. Guidelines can also be restrictive via prescriptive solutions to design problems, and sometimes different guidelines provide conflicting information. They also cannot offer any quantitative prediction of populations accommodated by the design (Dong, 2004; Goodman et al., 2007).

Empathic simulation is another method that can be used to assess products without the presence of actual users. Wearable simulators can generate empathy and even roughly estimate levels of disability in the population if calibrated (Cardoso, 2005). This method builds on designers' preference for hands-on methods, experiential learning and understanding of users' needs. However, these methods are limited in trying to reproduce the actual experience of living with the effects of ageing or a full time disability, and they currently cannot be applied to simulating cognitive capability loss (Cardoso, 2005). Though some simulators are configurable to simulate multiple capability loss, such as the EDC wearable capability simulator (Cardoso & Clarkson, 2006), they cannot realistically simulate all given types of loss. In any event, it would be a tedious task to attempt all combinations. Expert evaluation is also used in assessing a design for usability and accessibility (Cardoso, 2003; Nielsen, 1993). The method depends on the individual experience and expertise of the assessors and usually more than one assessor is required to capture most problems with the design.

Another approach to analytical evaluation is via the use of predictive models and computer simulation. Models of the user are used as the basis for predicting human performance parameters such as time taken and errors made. Various cognitive architectures and related methods exist such as GOMS, EPIC, ACT-R, and SOAR for basing the evaluation system (Kieras, 2003). These systems allow for a prediction of performance parameters such as the total time to perform a task. However, many of these systems address the functioning of 'average' users. For example, GOMS predicts time taken based on normative population values. The modelling technology has not yet developed to take account of ageing or disabled users in any major way (Kondraske, 2006a; Kondraske, 2006d).

Finally, analytical evaluation can be conducted using measured human capability data that is captured in a database and used to predict performance parameters and populations accommodated. They are based on principles originally developed in the field of anthropometrics (Bittner, 1979; Bittner & Moroney, 1975). These methods, which simulate "virtual user trials", take as inputs product design parameters such as dimensions, layout, forces etc. and task parameters such as actions required to reach a goal. Using these inputs, a designer is able to predict the level of accommodation for a given design in terms of the number of people excluded in a given population. Examples include HADRIAN (Porter et al., 2004) and the Exclusion Audit tool (Cardoso, 2005). This approach is valuable, as the use of this type of simulation gives the designer a sense of how the product should be improved (Meister, 1995). The research presented in this thesis aims to investigate supporting this type of predictive analytical evaluation. Specifically, an assessment is needed of the required human capability data, and the predictive abilities of such data. The next section looks at these predictive methods in more detail.

2.4.4 Predictive Methods for Inclusive Design Evaluation

There are few examples in the literature that attempt to utilise population representative user data for Inclusive Design evaluation. Clarkson et al. used population representative disability data from the 1996/97 Great Britain Disability Follow-Up Survey to quantify design exclusion (Clarkson et al., 2003b; Waller, Goodman-Deane, Langdon, Johnson, & Clarkson, 2009a; Waller et al., 2010b). Seven disability scales were used to evaluate products for demand on user capability, and estimates of proportions of the population that are excluded can be derived. Five of these scales were also used by Cardoso as the basis of an analytical evaluation tool (Cardoso, 2005). Similarly, Brinck used data from the US Census disability surveys in order to perform cost-benefit analysis of market size (Brinck, 2005). This involved quantifying the number of people excluded from current market segments and calculating the benefits of increased sales if they are included. Carlsson also showed the value of multivariate profiles of capability data in predicting device accessibility, as multiple capability losses tend to cluster into groups (Carlsson et al., 2002).

The data collected in large scale population surveys are not specifically designed to support product evaluations; however they are valuable to designers and manufacturers in developing a population perspective and supporting the business case for Inclusive Design (Dong, 2004). The key to these methods are that sensory, cognitive and motor capabilities need to be taken into account at the same time to look at how individual *people* (represented by capability profiles) are excluded from the product. Then, based on the sample in the capability database, generalisations are made about user *populations*. The following sections describe state of the art analytical evaluation systems with an eye toward deficiencies in the supporting capability database.

2.4.4.1 HADRIAN, SWORDS and SHIELDS

HADRIAN is a computer based evaluation tool developed at Loughborough University that performs 'virtual user trials' based on a database of user data (Marshall et al., 2010; Porter et al., 2004). This data set contains mostly anthropometric data, and also includes some capability data such as exertable forces. Though the sample in the database is relatively small, HADRIAN demonstrates the advantage of multivariate accommodation where multiple variables/characteristics of the user are considered at a time. A CAD model of the product is used as input to the HADRIAN progam. A task analysis is then entered into the software, where each action to be performed with the product is described with the appropriate action words and parameters. The software then evaluates the product against a person at a time and outputs an estimation of the proportion of the sample that is excluded. It is possible to output the exclusion for each step in the task sequence, or the cumulative exclusion across task steps. Screenshots of the HADRIAN interface are shown in Figure 2-6.

HADRIAN was extended by Goontekille (2003) who introduced optimisation and constraint modeling in a system called SWORDS/SHIELDS. It works by taking output from HADRIAN that represents the extent of mismatch per user (such as dimensions) and constructs objective functions to be minimised. Thus the result of the software operation is a list of design variables, in this case product dimensions, that exclude the minimum number of people in the HADRIAN database. This is one of the few examples of Inclusive Design optimization in the anthropometric domain.

Figure 2-6 Screenshots of HADRIAN utilising anthropometric data and calculating proportions excluded

Though HADRIAN demonstrates the value of multivariate capability profiles and evaluating across the sequence of use, the capability database is lacking in terms of sample size and the inclusion of other sensory, cognitive and motor capability variables that are important in everyday product usage. The database is biased toward anthropometrics, and thus it is limited as a general product evaluation tool.

2.4.4.2 Exclusion Audit and Capability Scales

Previous research (Keates & Clarkson, 2003a) has investigated the relationship between capabilities of the population at large and guidelines for the design of features of products. This research suggested that a good representation of the capability range of individuals can be made on a three axis scale derived from the basic psychological dimensions of sensory, motor and cognitive capability. To populate this model, there are many sources of capability data addressing sensory, cognitive or motor domains individually. However, recent research has focused on one of the most complete representative disability data sets; the 1996/97 Great Britain Disability Follow-up Survey (DFS) (Grundy, Ahlburg, Ali, Breeze, & Sloggett, 1999; Johnson et al., 2010; Waller et al., 2010b).

This survey was based on a measure of severity of disability established through the consensus of judges including medical experts and disabled people (Martin & Elliott, 1992). The DFS survey introduced thirteen capability scales of which seven are most relevant to product evaluation: locomotion, dexterity, reach and stretch, vision, hearing, communication, and intellectual functioning (Keates & Clarkson, 2003a). Figure 2-7 shows the seven scales representing the capability profile of a person with various scores along the scales. By utilising a linear combination model, an overall assessment of the severity of disability could be calculated as a score from 1 to 10 given the scores on each of the scales. Though the scales lack the granularity and completeness to evaluate all aspects of consumer product interaction, they provide a unique set of multivariate capability data that is representative of the Great Britain population.

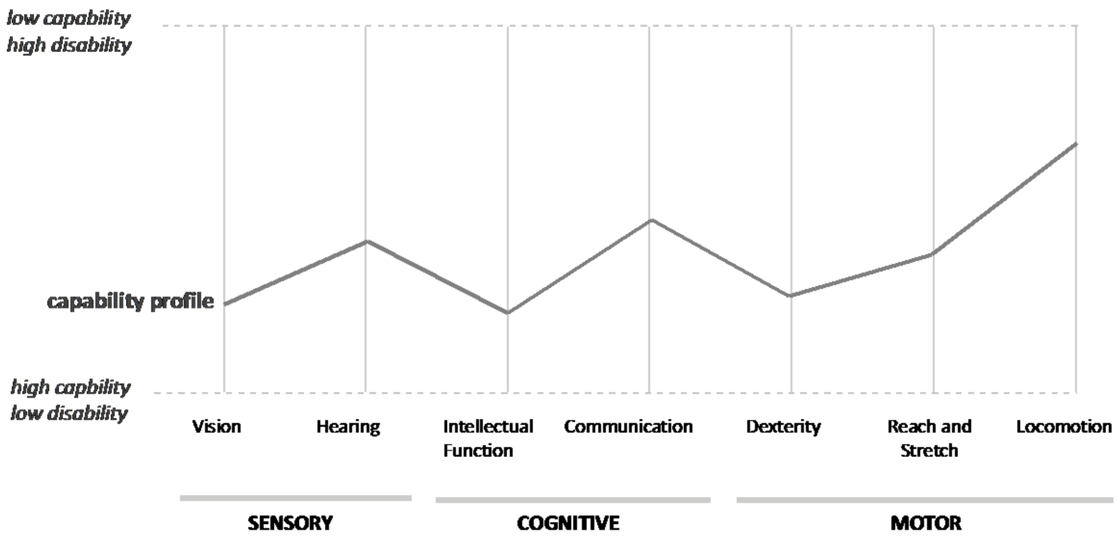


Figure 2-7 Illustration of the DFS Disability Scales

Though these scales were found to have advantages such as coherence, utility, and statistical validity for their intended purpose (Dong, 2002), they also presented certain disadvantages when adopted for evaluating product interaction. Firstly, the scales do not represent the discontinuous, non-linear nature of capability loss. They simplify the phenomena of multiple capability loss into a naive linear additive model, which might not represent what actually occurs. Secondly, the scale points are not specific enough to perform detailed evaluations of product interface features. Table 2-7 shows the vision disability scale as an example. This scale ranges from a score of 0.5 (high capability or low disability) to a score of 12 (low capability or high disability). The scale points are given in terms of the capability to perform everyday visual tasks, with some using the descriptor "having difficulty" and others using "cannot."

Table 2-7 Vision Disability Scale

Score	Description
12.0	***Cannot*** tell by the lights where the windows are
11.0	***Cannot*** see the shapes of furniture in a room
10.0	***Cannot*** see well enough to recognise a friend if close to his face
8.0	***Cannot*** see well enough to recognise a friend who is an arm's length away
5.5	***Cannot*** see well enough to read a newspaper headline
5.0	***Cannot*** see well enough to read a large print book
4.5	***Cannot*** see well enough to recognise a friend across a room
1.5	***Cannot*** see well enough to recognise a friend across a road
0.5	***Has difficulty*** seeing to read ordinary newspaper print

For example "large print book" and "newspaper headline" convey approximate size ranges of text, but these descriptions cannot be effectively used to evaluate a precise size of text. In addition, the scales cannot address the problem of evaluating a *small change* in text size. Thirdly, the scales confound a number of known underlying capabilities. For example, the visual task of object recognition in various lighting conditions is different to reading high contrast text, and as such both tasks demand the engagement of different combinations of visual functions. Finally, the scales were founded on judges' notions of disability and a self-report questionnaire rather than on objective measurement (Dong, 2002; Langdon, Japikes, Clarkson, & Wallace, 2003). This method could lead to inaccuracies in the data collected.

Table 2-1 shows the locomotion disability scale as another example. Here different actions such as walking, balancing and bending are all confounded in the scale together with 'can' and 'cannot' descriptors. Just as with the vision scale, the locomotion scale proves difficult to use for detailed product assessments.

Table 2-8 Locomotion Disability Scale

Score	Description
11.5	*Cannot* walk at all
9.5	***Can*** only walk a few steps without stopping or severe discomfort. ***Cannot*** walk up and down one step
7.5	Has fallen 12 or more times in the last year
7.0	Always needs to hold on to something to keep balance
6.5	***Cannot*** *walk* up and down a flight of 12 stairs
5.5	***Cannot*** *walk* 50 yards without stopping or severe discomfort
4.5	***Cannot*** *bend* down far enough to touch knees and straighten up again
4.0	***Cannot*** *bend* down and pick up something from the floor and straighten up again
3.0	***Cannot*** *walk* 200 yards without stopping or severe discomfort. ***Can*** only walk up and down a flight of 12 stairs if holds on and takes a rest. Often needs to hold on to something to keep balance. Has *fallen* 3 or more times in the last year.
2.5	***Can*** only walk up and down a flight of twelve stairs if holds on (doesn't need a rest)
2.0	***Cannot*** bend down to sweep up something from the floor and straighten up again
1.5	***Cannot*** walk up and down a flight of stairs if goes sideways or one step at a time
0.5	***Cannot*** walk 400 yards without stopping or severe discomfort

It is evident therefore, that the scales are not at a hierarchical level that is suitable to carry out a detailed product evaluation, and there may be problems with inter-rater reliability (Johnson et al., 2010; Waller et al., 2010b). Given these limitations, design exclusion audit tools based on this dataset (Figure 2-8) have proven useful for making approximate predictions of excluded populations (Cambridge Engineering Design Centre, 2011; Cardoso, 2005; Waller et al., 2010b). In addition, examining the exclusion across each action in the task sequence is also valuable to pinpoint areas of maximum exclusion (Waller et al., 2010b). However, as the dataset becomes dated and the limitations in the scales become more apparent through usage, there is a need to re-examine the nature and composition of capability data to support analytical evaluation tools.

Figure 2-8 Exclusion audit tools showing interaction maps and an exclusion calculator based on the DFS capability scales

2.4.5 *Advantages and Disadvantages of Current Methods*

In the preceding discussion, state of the art evaluation methods were shown to utilize the concepts of multivariate capability profiles, exclusion across the task sequence and product evaluation supported with databases of user capability data. The idea of multivariate accommodation is a fundamentally important idea – it is advantageous to look at capability profiles of each person in addition to looking at one capability variable across many people (Carlsson et al., 2002; Marshall et al., 2010; Waller et al., 2010b).

These methods, however, also have certain disadvantages. Current capability databases do not address the cognitive domain adequately for product evaluation (Johnson et al., 2010; Waller et al., 2009a). This important dimension of interaction needs to be considered together with sensory and motor capabilities in order to perform a comprehensive product evaluation. In addition, current capability databases are not adequate in terms of their comprehensiveness and detail for conducting a full product audit with respect to the scope of actions that users take in activities of daily living (Johnson et al., 2010).

2.5 The Capability Data Problem

In the first instance, human capability data to support Inclusive Design is fragmented, lacking and mainly exists as small sample studies in different domains (Gyi et al., 2004). Population data sources such as the 1996/97 Great Britain Disability Follow-up Survey (DFS) (Grundy et al., 1999) exist, but are limited by the approach to measurement described in the preceding section (Johnson et al., 2010). Ageing and medically oriented data sources exist that can shed some light on capability distributions in the population, for example the English Longitudinal Study on Ageing (ELSA) (NATCEN, 2008), the Medical Research Council Cognitive Function and Ageing Study (MRC, 2008), Hearing in Adults (Davis, 1994) and the RNIB Vision Survey (RNIB, 1991). However, these studies are not geared specifically toward providing designers with capability data that can be used in product assessment (Johnson et al., 2010).

Human Factors and Ergonomics capability data sources exist including anthropometric and strength data (Feeney, 2002; Smith et al., 2000), and other studies supporting the use of products by disabled people (DTI, 2000; Kanis, 1993; Norris & Wilson, 1997). Some data books such as OlderAdultdata (Smith et al., 2000) aim to solve the problem of disparate data by collating measures from different studies and populations. This interim approach is useful till a comprehensive, integrated capability data set is available. In one of the few examples of the required integrated approach to capability measurement, Steenbekkers et al. captured an integrated, multivariate data set for 750 Dutch users (Steenbekkers & VanBeijsterveldt, 1998; Steenbekkers et al., 1999). The data is comprised of 79 physical, sensory, cognitive and psychomotor capability variables that are thought to be relevant to product interaction. There is no equivalent data source for the UK population. Even so, the data captured by Steenbekkers et al. was published in the traditional book format with tables of data, thus reducing its chances of being used by designers (Wilcox, 2007).

Another problem involves capability data quickly becoming outdated due to changing demographics (United Nations, 2009; Yoxall et al., 2006). It is imperative to develop standardised measures that are administered regularly to accurately track the changing capability structure of the population at large.

2.5.1 Accuracy of Predictions Using Capability Data

Even more fundamental to the lack of design relevant capability data is the problem of the validity of predictions made from such data. In other words, within such data sources, are the right variables being measured? Research has shown that predictions of real-world interaction problems in disabled populations produce variable results (Kanis, 1993; Steenbekkers & VanBeijsterveldt, 1998; Steenbekkers et al., 1999). In a study of control operation by physically disabled users, Kanis found that he was able to accurately predict users' difficulty in operating controls for a little more than 50% of the cases (after measuring maximum force exertions and the force required by the control). For a third of the cases, the difficulty was overestimated, and for one tenth of the cases, the difficulty was underestimated (Kanis, 1993). Steenbeekkers et al. also concluded that laboratory measures have limited predictive value for difficulties experienced in daily life (Steenbekkers & VanBeijsterveldt, 1998). They also mention that it is not clear how individual measures of capability combine to enable an individual to successfully complete a task. Confounding factors include the various compensation, adaptation and coping strategies employed together with the interaction of sensory, cognitive and motor capabilities (Steenbekkers & VanBeijsterveldt, 1998).

Given this situation, there is scope for further research into understanding how various capability measures of disabled populations interact and relate to the real world performance of tasks with consumer products (Johnson et al., 2010). If models could be developed that match product demand measures to user capability measures with success at predicting real world difficulty, there will be a valid foundation for data collection to support inclusive analytical evaluation methods. There is therefore a need to determine the underlying models of interaction, which will eventually focus user capability data collection efforts for design applications.

2.6 Summary of Research Problem

Based on the preceding sections, two main areas requiring further research are identified. Firstly, a need continues to exist for product evaluation methods that can support the designer in evaluating consumer products for Inclusive Design (Cardoso, 2005; Waller et al., 2010b). Specifically, evaluation methods are required that can be used prior to prototyping or testing with actual users (Vanderheiden & Tobias, 1999). Evaluation methods are also required that can help estimate design exclusion on the group and population level together with an understanding as to *why* particular features are problematical (Cardoso, 2005; Dong, 2004).

Secondly, there is a lack of coherent, comprehensive and up-to-date user data to support such inclusive evaluation methods (Dong, 2004; Johnson et al., 2010).

Based on an assessment of the literature in the field of Inclusive Design evaluation, it was found that there continues to be a lack of understanding of the theoretical aspects of Inclusive Design and a lack of guidance from research on how to evaluate inclusive product designs (Warburton, 2003). An argument was made for the usefulness of analytical methods for inclusive product evaluation, where a population level approach is needed when considering users with reduced sensory, cognitive and motor functional capability (Carlsson et al., 2002). Only population representative user data can provide accurate estimates of excluded user groups. Not only must user capability data be representative of the population at large, but it must also be design relevant and multivariate, consisting of sensory, cognitive and motor capability data all in the same database (multivariate data). Most current evaluation methods are not based on the integration of sensory, cognitive and motor capability data at a suitable level of detail.

Most human factors data tends to be for relatively homogenous populations (Kondraske, 2006a; Kondraske, 2006d). In addition, using capability data to make real world predictions of difficulty and exclusion for disabled people is not well understood (Kanis, 1993; Kondraske, 2006a; Steenbekkers & VanBeijsterveldt, 1998). As a precursor to further developing analytical evaluation approaches and collecting human capability data to support these approaches, this more fundamental problem needs to be addressed.

Indeed the lack of adequate data on disabled populations and the related issue of regulations that govern the testing protocols of products could be considered unethical (Bix, de la Fuente, Pimple, & Kou, 2009). Bix provides the example of test protocols for child resistant medical packaging requiring test subjects who do not have 'obvious or overt physical or mental disabilities.' She argues that this nebulous condition should be removed from the test protocols as it constitutes a violation of the biomedical ethical principle of justice (Bix et al., 2009). It also excludes many users of the target population who are likely to utilise various forms of medication. With the pressures of increasing legislation to avoid discrimination against various groups in society, amendments and additions should be made to the law requiring the consideration of older and disabled users in all aspects of life.

The research presented in this thesis therefore aims to address this fundamental problem by theoretically and empirically investigating the relationships between product demands and user capability measures in the context of a capability-demand product interaction framework.

By taking a holistic view, the interaction of user capabilities will also be investigated, so that a better understanding of user-product interaction can be derived. This should help to determine what type of capability-demand model and user capability data could form the basis of valid and robust analytical evaluation methods.

2.7 Chapter Summary

This chapter examined the literature on Inclusive Design, and was structured around three key issues: (1) the nature of Inclusive Design problems and supporting the designer in evaluation and decision making, (2) knowledge of users and their characteristics, and (3) product evaluation methods and supporting data for analytical Inclusive Design evaluation. Current tools and methods for Inclusive Design evaluation were reviewed and their advantages and limitations were discussed. After arguing for the need for analytical evaluation approaches based on population representative capability data, the review culminated with a clear need for fundamental research into understanding how user capability data can be used to predict problems in real-world interaction. This knowledge can then be used to form the basis of analytical evaluation approaches for Inclusive Design.

Chapter 3 **Research Approach**

3.1 Chapter Overview

This chapter presents the research approach employed to address the research questions. The research progressed in three phases of reviews, exploring inclusive interaction and experimental testing. The research methodology encompassed a mixed method approach, including both quantitative and qualitative methods. The rationale for the methodology selected will be described and critiqued.

3.2 Methodological Approach

The search for an adequate methodology originated from the three main research questions, which involved exploring the adequacy of a *user capability-product demand interaction model* as a useful model for Inclusive Design evaluation. As described in Chapter 2, previous work demonstrated a lack of analytical evaluation methods that could predict accessibility problems and design 'exclusion' at the group or population level. More fundamentally however, it was shown that the underlying theory and associated models of user capability and product demand required further investigation before such analytical methods could be developed. Specifically, measures of human sensory, cognitive and motor capabilities must demonstrate accurate predictions of problems in real-world product interaction if these measures are to be used for analytical evaluation. The three main exploratory research questions addressed are described as follows:

R1: What theoretical models exist for understanding the relationship between human functional capabilities and real world task performance in ageing/disabled populations?

Given that the background literature review revealed a clear need for further understanding the theoretical aspects of Inclusive Design assessment (Kondraske, 2006c; Steenbekkers et al., 1999; Warburton, 2003), the first exploratory research question seeks to understand what extant models are available for modelling the interaction of human capabilities in the performance of tasks in older and disabled populations. In addition, the question implicitly asks about the comparative performance of these extant models. The nature of the question implies methods of systematic literature review in published studies of ergonomics, interaction and biomedical models of human function.

R2: What are the key elements of human functional capability that influence inclusive product interaction?

The second research question deals with the ontological foundations of the theoretical models, i.e. it seeks to extract the domains, hierarchy and elements of human function that are known to be influenced by ageing/disability and at the same time to be essential for product interaction. A systematic literature review of published literature is also implied by this research question. In addition, the question is amenable to investigation by empirical/experimental testing and validation to determine if indeed selected human functional capability measures factor in the performance of tasks with products.

R3: What relationships exist between measures of human functional capability and measures of task performance in the context of a user capability-product demand model of interaction?

The third research question deals specifically with investigating the relationship between measures of user capability, product demand and task performance in inclusive interaction. Cognisant of extant theories and models for explaining these relationships as a result of asking research questions one and two, an experimental or 'testing-out' method is implied.

In choosing a research methodology, the most important consideration is selecting a framework that can answer the individual research questions (Robson, 2002). Robson (2002) presents a classification based on the purpose of the enquiry consisting of exploratory, descriptive, explanatory and emancipatory research. These categories are not mutually exclusive, and a research project can span different categories. Robson also suggests that multiple methods can be used to address different but complimentary research questions, and different methods may be appropriate at different stages in the research process (Robson, 2002, p. 371). Leonard (2006) also provides a tripartite classification of descriptive, experimental and evaluation research in the field of Human Factors and Ergonomics. Evaluation research can comprise elements of descriptive and experimental research as it aims to describe the people using a system, to understand the effects of human-system interactions and to evaluate methodologies and measurement tools (Leonard et al., 2006, p. 312).

Considering these research strategies, and the *'what'* question element of the three closely linked research questions, the methodology adopted is essentially *evaluative* in nature, utilising multiple methods (both quantitative and qualitative) at different stages in the research process. Thus, the research questions are addressed by utilising various methods including literature reviews, expert consultations, secondary data analysis, product analysis/inspection, user observation, and empirical/experimental user testing.

3.3 Research Process Overview

Table 3-1 gives an overview of the three phase research process consisting of (1) reviews, (2) exploring inclusive interaction, and (3) experiment. Each phase will be described in the following sections.

Table 3-1 Research Methods applied at different stages of the research process

Research Phase	Methods Applied	Thesis Chapters
Phase 1: Reviews (Addresses R1 and R2)	Review of Theoretical Foundations: Literature Review, Framework Development Review of User Capabilities: Literature Review, Expert Interviews Review of Existing Capability Data Sources: Literature Review, Expert Interviews, Secondary Data Analysis of ONS DFS Data.	4 5 2, 5
Phase 2: Exploring Inclusive Interaction *(Addresses R2 and R3)*	User observation with toasters (7 Participants with different capability profiles) Product demand analysis	6
Phase 3: Experiment (Addresses R3)	Experimental study to explore relationships between user capabilities, product demands and task performance 4 consumer products tested: clock radio, mobile phone, blender, and vacuum cleaner 19 Participants with different capability profiles	7
Analysis and Recommendations	Recommendations on avenues for further research and development	8

3.4 Phase 1: Reviews

The first phase of the research process involved conducting systematic reviews to answer research questions one and two. Literature reviews, expert consultations, product analysis and secondary data analysis were used to better understand the theoretical basis for Inclusive Design evaluation and human capability measurement.

3.4.1 Review of theoretical foundations

The theoretical foundations of human-product interaction and product assessment methods were reviewed by surveying the Ergonomics/Human Factors literature. Similar concepts to the multidimensional capability space of the Inclusive Design Cube (Keates & Clarkson, 2003b) was found in the Elemental Resource Theory (Kondraske, 2006a), and these theoretical approaches were compared and contrasted (Chapter 4). In addition, a three phase generic framework for analytical product assessment was developed by integrating principles of analytical evaluation methods based on user data. Chapter 4 contains the full description of theoretical models and the analytical evaluation framework.

3.4.2 Review of user capabilities

The first phase of the research also involved systematically reviewing existing literature on changes in the sensory, cognitive and motor functional capabilities of ageing and disabled users. This entailed searching books, guidelines, journal articles and websites with Inclusive/Universal Design guidance to extract the common set of user capabilities presented.

The process entailed starting with the categorisation of the World Health Organisation's International Classification of Functioning, Disability and Health (World Health Organisation, 2001) and modifying the list accordingly with each new publication reviewed. Throughout the research process, experts and practitioners in the fields of vision (4 experts), hearing (2 experts), ergonomics (3 experts), occupational therapy (2 experts), medicine (2 general practitioners) and disability (3 experts) were consulted. These consisted of researchers at The University of Cambridge and other Universities, clinicians at Addenbrooke's Hospital and Anglia Ruskin University, and service staff at charities such as CAMSIGHT, University of the Third Age and the Papworth Trust. Data from interviews and visits were recorded using a combination of note taking, audio recording, digital photographs and video.

Researchers and experts provided input on the list of user capabilities, references to existing capability data sources and contacts for study participants. The synthesis of these reviews is presented in Chapter 4. The resulting information was also used to develop designer guidance which was published as an Inclusive Design Toolkit (Cambridge Engineering Design Centre, 2011; Clarkson et al., 2007).

3.4.3 *Review of product demands*

The research project was restricted to investigate every-day consumer products that support activities of daily living (ADL) and Instrumental Activities of Daily Living (IADL) that older and disabled people are required to use in daily life. Desktop computer software and websites were not specifically considered. Consumer product interfaces were analysed by looking at product catalogues and brochures in print and on the web, and visiting department stores to informally observe people buying and testing products. Assistive devices were also investigated to understand how general consumer product design could be improved to cater to users with reduced functional capacity. Figure 3-1 shows product demonstrations at CAMSIGHT, a charity for people with visual disabilities. Data was recorded using a combination of note taking, audio recording, digital photographs and video.

Figure 3-1 Product demonstrations at CAMSIGHT, a charity for people with visual disabilities in Cambridge

3.4.4 *Review of Existing Capability Data Sources*

As part of the review process, data sources on user capabilities and their potential for application to product assessment were evaluated. In addition to website searches and references from domain experts, institutions such as the Office for National Statistics, The Royal National Institute of Blind People (RNIB) and The Royal National Institute for Deaf People (RNID) were contacted for current data sources on user characteristics. References to all data sources were compiled and recorded using Thomson Reuters' Endnote bibliographic management software.

Integrating different data sources was not possible due to different disability estimates and definitions (Bajekal et al., 2004; World Health Organisation, 2011). Using a statistical meta-analytical approach for different data sources was considered, but this was abandoned given that such a method would not yield a sufficiently integrated data set from which estimates of exclusion could be made. In addition, it would have been difficult to access original datasets from different institutions.

It was decided that the 1996/97 Disability Follow-Up Survey (Martin, Meltzer, & Elliot, 1988) provided the best multivariate data set for investigating capability relationships that reflects population level disability. To understand the distribution and grouping of disability in the UK population, secondary data analysis was performed on the 1997/98 Disability Follow-Up Survey (DFS) data using SPSS analysis software. Using 13 disability scales and 15 categories of medical conditions, the data was explored using frequency plots and exploratory cluster analysis (Everitt, Landau, & Leese, 2001). Two medical experts assisted with the interpretation of the cluster analysis by examining the visual clustering outputs (dendrograms) and by providing reasons for the structure in the data. The details and results of this study are presented in Chapter 5.

3.5 Phase 2: Exploring Inclusive Interaction

In the second phase of the research process, an observational study was conducted with seven participants using two toasters. The reason for this was twofold. Firstly, the study aimed at getting a better understanding of the intricacies of inclusive interaction which would serve to validate the literature and inform assumptions made for analytical methods. Secondly, it was necessary for the researcher to develop experience working with older and disabled people, and observing their experiences first hand. Working with older and disabled people requires the development of special skills and understanding in an empathic model, and only through direct experience this becomes possible (Cassim & Dong, 2003).

The study was designed so that a relatively simple toaster and a relatively complex toaster could be compared. This translated to two products of relatively low demand and high demand. A product demand analysis of the simple toaster was also carried out utilising task analysis and state charts to encapsulate the demanded mental model of operating the device. The study did not strive for statistical generalisation, thus participants were sampled using convenience (heterogeneous) sampling (Robson, 2002, p. 265). The details and results of this study are presented in Chapter 6.

3.6 Phase 3: Experiment

In the third phase of research, an experimental study was conducted with 19 older and/or disabled users performing tasks with four consumer products (a clock radio, a mobile phone, a blender and a vacuum cleaner). The study aimed to investigate the relationships between user capability measures, product demands and task performance outcomes. The study design involved first measuring selected sensory, cognitive and motor capabilities of users before they proceeded to use the products in randomised order. The users were videotaped while using the products and they rated a range of selected actions for difficulty using a visual difficulty rating scale. This method allowed for objective (time and errors) and subjective (difficulty) measures to be collected for each user, thus allowing for direct comparisons of user capability measures, product demand measures and task performance outcome measures via co-relational studies.

The capability testing equipment for this study was acquired using a combination of equipment purchases from vendors in the UK and the USA, together with manufacturing handles and mounting fixtures with the aid of another researcher in the Cambridge Engineering Design Centre. Computer based vision (Test Chart Pro) and cognitive testing software (CANTABeclipse) were also researched and purchased to form the suite of capability tests that were administered to users. The capability testing kit required time and effort to put together, as different parts had to be sourced from different vendors with different lead times. Once all the equipment was in place, three pilot studies were run to improve the study method, and the study took place over three weeks.

After receiving ethical approval from The University of Cambridge's Ethical Committee, users for the study were recruited based on advertisements and calls to various Cambridge charities. Once the experimental study was completed, the data was analysed and the resulting relationships are presented and discussed in Chapter 7.

3.7 Critique of the research methodology

Various decisions were taken in the overall research process, and these are discussed in this section. Firstly, the scope of the research was limited to measurable sensory, cognitive and motor capabilities. Issues such as motivation and aesthetic/emotional response are important and play a role in the perception of usability and difficulty when using a product. However, these were not considered in order to constrain the research project to only functional capabilities.

A persistent problem remains with sampling users for Inclusive Design studies. On the one hand, random sampling might not capture 'edge' cases i.e. users with some severe form of disability. In addition, the population from which users are being sampled is very heterogeneous, and trying to capture a 'representative' sample would result in a very large sample size to cover all cases. On the other hand, small sample sizes can be biased by users with some extreme forms of disability, leading to solutions that might only satisfy a small group of such users (Kanis & Arisz, 2000).

The approach taken in this research was to investigate the interaction phenomena at hand using relatively small samples so that the nuances of interaction could be observed in detail. The main argument is that a capability-demand model with its assumptions should be examined for individual users across the disability spectrum. Following this, a larger scale investigation for model validation would be necessary. However, studies with larger sample sizes across the disability spectrum require more research resources, time, effort and funding together with increased logistical difficulties. However, as an applied empirically driven discipline, ergonomics and human factors research relies on such validation. The generic analytical evaluation framework developed in Chapter 4 should be based on a capability database comprising measures representative of the larger disabled population. The statistical problems of determining the size of such a database and generating estimates of exclusion to the wider population remains a problem that is beyond the scope of this research.

The aim of the experimental study was to look at user capabilities, product demands and task outcomes to see what relationships exist. The approach has been to start with the most common model of user capability-product demand interaction (linear co-relational models) and look at its ability to account for the relationships observed. The experimental study presented in Chapter 7 serves falls within the first and second stages in the research agenda set out in Chapter 4. The methodology developed for the experimental study can be adopted for larger scale studies to investigate such inclusive interaction phenomena.

3.8 Chapter Summary

This chapter described the three stage research methodology adopted to address the three research questions together with a description of the methods used in the research process. The detailed results and analyses are given in Chapters 4 & 5 for Phase 1, Chapter 6 for Phase 2 and Chapter 7 for Phase 3.

Chapter 4 **Theoretical Considerations**

4.1 Chapter Overview

In this chapter, relevant theoretical frameworks are reviewed. Firstly, user capability-product demand theory as the basis for applied ergonomics and human factors is described. Secondly, theoretical models for the interaction between human functions and tasks (including the relationships between low-level and high-level functions) are reviewed. Finally, the chapter ends with a description of a generic framework for analytical Inclusive Design evaluation and the requirements for operationalising the framework.

4.2 Interaction: User Capability-Product Demand Theory

The ideas of user capability and product demand provide a useful framework for design evaluation where the sensory, cognitive and motor demands made by a product are compared to the capability levels of the target user population (Bridger, 2003; Clarkson & Keates, 2003b). This theory of user-product compatibility is well known and forms the basis of the field of ergonomics and human factors (Bridger, 2003; Karwowski, 2002). These concepts are shown diagrammatically in Figure 4-1.

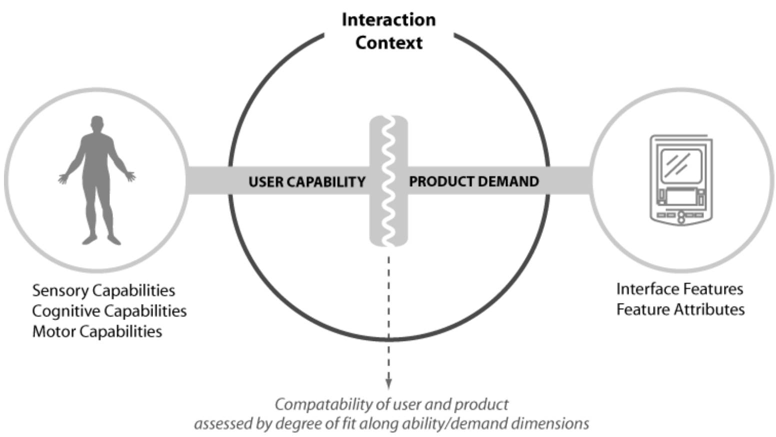


Figure 4-1 An illustration of the relationship between user sensory, cognitive and physical capabilities and the demands made on the user by the product

When a user interacts with a product, there is a cyclic process of perception, cognition and action through time in a given physical and social context, as shown in Figure 4-2 (Monk, 1998; Norman, 2002). The effects of previous actions are first perceived and interpreted according to expectations. These interpretations are then evaluated in terms of the user's previous intentions and current goals. Based on this comparison, an intention is formed, and this intention is further operationalised into a sequence of actions that can satisfy the intention. This mental plan is then executed on the product, which changes its state. At this point, the cycle repeats itself. In this way, a user moves through successive cycles from a starting state to an end state where hopefully a goal will be achieved.

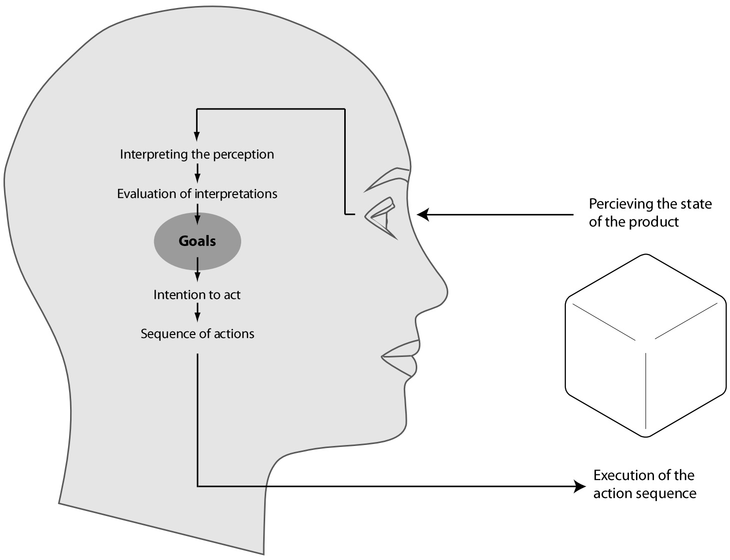
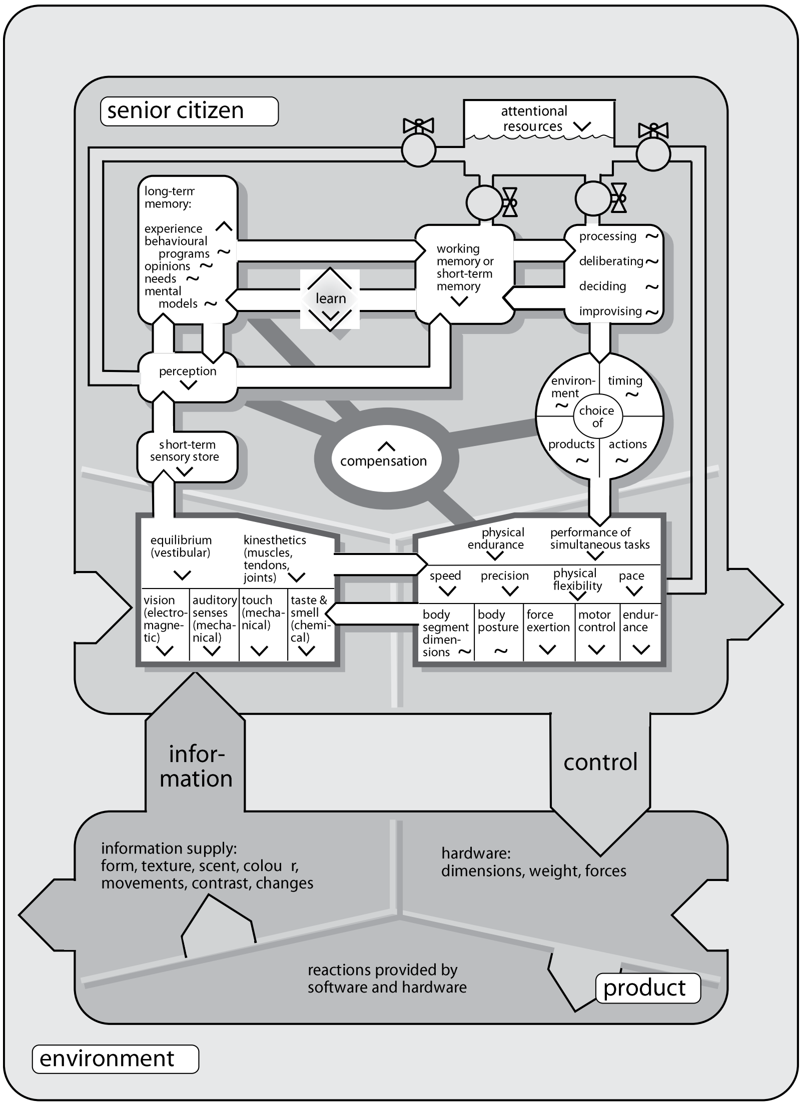


Figure 4-2 The cycle of interaction: perception, cognition and action after (Norman, 2002)

The stages of perception rely on human sensory capabilities such as vision, hearing, taste, smell and touch. The stages of interpretation, evaluation, intention to act and sequencing of actions all rely on human cognitive capabilities such as working memory, long-term-memory and the processing capabilities of the human mind. Execution of the planned action sequence relies on human motor capabilities such as speech, motor control and strength in the limbs to physically perform actions in the world. Such frameworks appear in the literature embodying similar concepts with different levels of detail.

A particularly comprehensive framework for product interaction was developed by Freudenthal (1999). It includes the directionality of changes in sensory, cognitive and motor capabilities with age (Figure 4-3). As evident from the diagram, the majority of capacities tend to decrease with increasing age. The components and scientific data to back up such frameworks are drawn from diverse fields including cognitive and social psychology, ageing studies, disability studies, biomedical sciences and ergonomics research. As the theoretical knowledge base is built up from these constituent disciplines, frameworks for the architecture and function of the human element in the interaction system evolve to integrate new findings. Therefore, frameworks such as the one presented in Figure 4-3 can be considered to be evolving meta-frameworks with interacting sub-systems.



A conceptual framework of senior-product interaction.
V a decrease in capacity that easily can influence product use negatively;
~ a change in processes or capacities that can influence product use in some way;
Λ an increase in capacity that can improve product use if the conditions are right.

Figure 4-3 A Framework for Senior Product Interaction (Freudenthal, 1999)

On the product side of the interaction equation, demand levels are set by the attributes of the product interface features. For example, a textual display on the product chassis will be designed with a certain text size, font style and colour, all placed on the product surface of another colour and material finish. This combination of design attributes sets a level of visual demand on the user. In a similar way, combinations of other attributes lead to cognitive and motor demands. Therefore, using a capability-demand framework as presented in Figure 4-1 is a useful starting point in considering analytical evaluation methods for Inclusive Design. By focusing on ways to measure product demands and relate them to measures of user capabilities, an estimate of the compatibility or fit between user and product could be established.

Even so, there is an important element to the interaction framework that requires explicit attention. User capabilities and product demands are always linked by the *task* that is to be performed. Carroll (1993) defines capability as follows: *"As used to describe an attribute of individuals, ability refers to the possible variations over individuals in the* ***liminal*** *levels of task difficulty (or in derived measurements based on such liminal levels) at which, on any given occasion in which conditions appear favourable, individuals perform successfully on a* ***defined class of tasks****"*. Carroll further defines a task as: *"... any activity in which a person engages, given an appropriate setting, in order to achieve a specifiable class of objectives, final results, or terminal states of affairs."* Thus a cognitive task and ability is defined as: *"... any task in which correct or appropriate processing of mental information is critical to successful performance. A cognitive ability is any ability that concerns some class of cognitive tasks, so defined."* Following this definition, a sensory and motor capability can be defined similarly as an ability that concerns sensory and motor tasks (Carroll, 1993; Fleishman & Quaintance, 1984; Kondraske, 2006a).

In other words, from these definitions, capability could only be understood when taking into account the task being performed *and* the level of performance required for success in that task. The level of performance required is set by the parameters of the task, which, in keeping with the capability-demand model, will be the product demand. The main point is that the *system* of user, product, task and environment and their interactions must all be considered. Capability variation is the variation in the threshold levels of performance among different people in the population of interest. In addition, capabilities and tasks are hierarchical in nature i.e. capabilities and tasks could be broken down into sub-capabilities and sub-tasks. Therefore, in operationalising user capabilities for measurement, a level of granularity must be decided. This issue is explored further in the next section on modelling human functional capability.

4.3 Modelling: Human Functional Capability

In order to characterise users based on functional capability, a deeper understanding of human function is required. Considering the human being as a system, there are various sub-systems that fulfil various functions. Figure 4-4 shows a useful framework for understanding the key relationship between functions of sub-systems (such as sensory, cognitive and motor systems) and functional capabilities of the person as a whole (Colenbrander, 2003). This diagrammatic adaptation from Colenbrander (2003), using the domain of vision as an example, highlights the links from disease to changes in the structure and function at the organ level (visual functions) and finally the effects of these changes at the person level (functional vision). The distinction between lower-level functions at the organ or human sub-system level, and functional abilities at the person level, is an important one for understanding the relationship between capabilities and tasks.

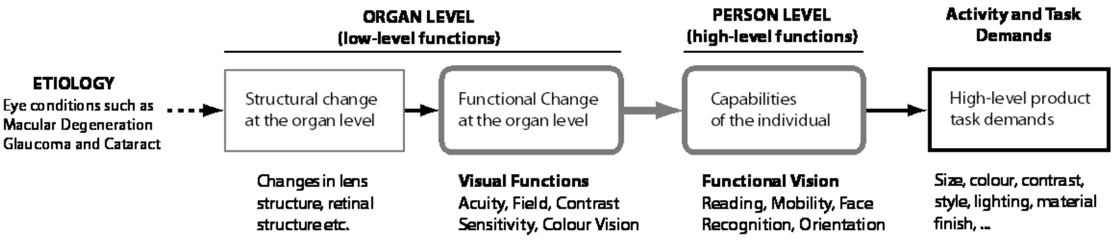


Figure 4-4 Illustration of the links between disease, organ structure and function and person function in the real world

Any analytical evaluation system should attempt to highlight problems that users would encounter if they interacted with a given product in the real world. As such, the primary interest then is human performance at the *person level*. For example, suppose one is interested in predicting if a given user can read a textual display (functional vision) on a product interface. If that user cannot perform the task, the next step would be to determine the changes to the design attributes (such as size, style, contrast, orientation, colour etc.) that would reduce the visual demand. The major question that needs to be asked is: what level of measurement is suitable for product evaluation - the organ level (e.g. low-level visual functions) or the person level (e.g. functional vision)?

To answer this question, the relationship between human sub-system function and person function in high level tasks requires examination in some more detail. Figure 4-5 shows examples of this relationship for the sensory, cognitive and motor domains. It is evident that the high-level functions are dependent in some way on the low-level functions. For example, the ability to grasp and lift an object depends on strength, control and range of motion of the hand and arm.

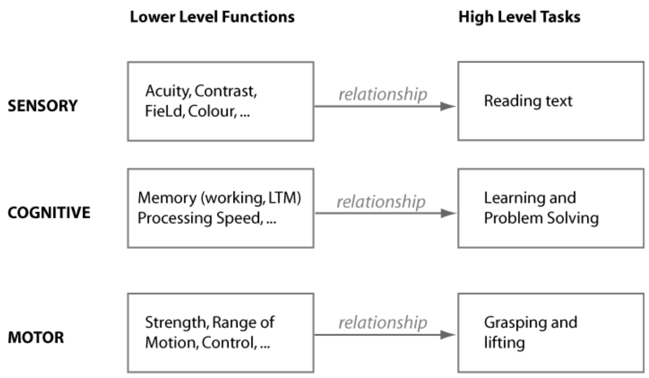


Figure 4-5 The relationships between lower level functions of human sub-systems and the performance of high level tasks in the sensory, cognitive and motor domains

If, on one hand, measures of high-level task performance are examined with functional assessment scales (rating scales), they would give estimates of the overall level of functioning of a person in a particular domain e.g vision or mobility. However, such scales are constructed by mixing different functions and combinations within the domains of performance (Kondraske, 2006d), and they aim only to assess *overall levels* of functioning. Thus by composing an overall measure of high-level task performance, it becomes difficult to isolate the contributions of lower-level functions, which are directly affected by product design attributes. In addition, because functional assessment scales address function at a global level using a set of high-level tasks, they are not generalisable in evaluating other high level tasks under different conditions.

Conversely, measures of human sub-system functional capacity (e.g. visual functions) could be used as indicators of high level task performance. So if predictions of real-world performance are required for real-world tasks, measures of capability at the human sub-system level are required. These capability measures would require a supporting model of how they should be combined to predict higher level task performance. In essence, a predictive model relating low-levels functions to high-level functions must be available.

4.3.1 Modelling the Relationship Between Hierarchical Levels

Engineering models of human performance can be useful in the design evaluation process by providing a simplified representation of a system that can be used to make predictions about human behaviour (John & Kieras, 1994; Meister, 1985; Zachary, Campbell, Laughery, Glenn, & Cannon-Bowers, 2001). However, there appears to be a relatively small number of extant models to adequately describe the relationship between human sub-system performance and person level performance (Kondraske, 2006a). Two of the most popular models will be described: (1) linear combination models and (2) non-linear resource economic models.

4.3.2 Linear Combination Models

The work of Fleishman in developing taxonomies of human ability requirements is relatively well known (Fleishman & Quaintance, 1984; Kondraske, 2006a), and the taxonomy of abilities can be useful at the conceptual level as described by Jacko (2001) (Jacko & Vitense, 2001). The predictive models developed by Fleishman utilise statistical regression models, which are linear combination models of low-level abilities used to predict high-level performance. However this type of model requires homogeneity among participants (i.e. they do not include participants with impairments) and it was found to have marginal performance (Kondraske, 2006a).

4.3.3 The Elemental Resource Model

The Elemental Resources Model (ERM) developed by Kondraske (Kondraske, 2006a) was developed to address the problems with linear models in predicting human performance. The ERM is a *performance based* model built on the more general concepts of General Systems Performance Theory (GSPT). The major elements of the theory relevant to this discussion will be presented here, while the reader is directed to (Kondraske, 2006a) for further details.

The basic concept behind the ERM is that human-sub systems can be considered to provide a finite set of *basic elements of performance resources (BEPs)* that are drawn upon when performing higher level tasks. As Figure 4-6 shows, the ERM is based on the concept of Mondaology or the existence of a finite set of such elemental units that combine into the complexity that is seen in the world. The level of performance on a given high-level task is therefore determined by the *limiting* lower-level BEPs i.e. humans performing high-level tasks may utilise many of these lower-level resources, but any one of these resources could be a limiting factor in the performance of the high-level task. This results in a *resource economic* view of human performance.

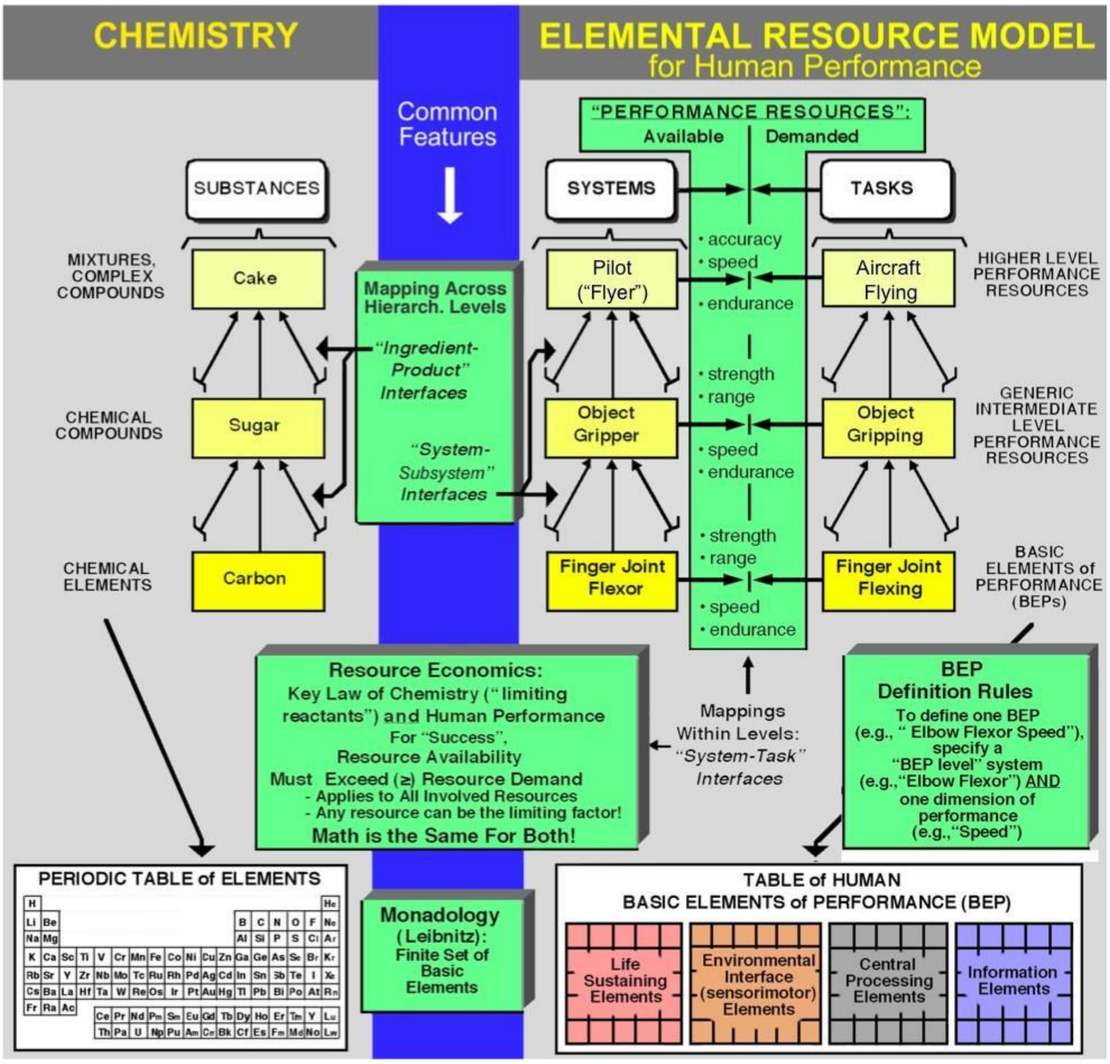


Figure 4-6 ERM model and the concept of Monadology (used with permission)

The following principles of the ERM are worth noting (Kondraske, 2006a):

The relationship between structure, function and performance: The distinctions between structure, function and performance are important. A system's structure enables its function, and is defined by its purpose or what the system exists to *do*. Systems can perform more than one function, though it is likely that only one function will be executed at a given time. Performance is a measure of *how well* a given system can execute its function. For example, it can be measured in terms of *strength*, *speed* or any other measure of capacity that can limit the system in task performance. The human system can be decomposed into a hierarchical structure down to a level where individual functional units serve a single function (e.g. elbow flexor: structure, elbow flexion: function). Each functional unit is characterised by a multidimensional performance space (Kondraske, 1988). Therefore, a basic element of performance (BEP) is defined by a functional unit and a dimension of performance e.g. visual memory capacity. By definition, a performance resource always starts at 0 and increases. The entire set of BEPs across all human functional units forms a finite resource pool, or a *multidimensional performance envelope*, from which an individual draws upon when executing a higher level task. The amount of performance resource drawn is dependent on the task to be performed, and task success results when all required resources are available in sufficient amounts.

Resource *availability* versus resource *utilisation*: There is also an important distinction to be made between resource *availability* and resource *utilisation*. Resource *availability* (such as maximum visual acuity, contrast sensitivity, locus of visual field etc.) could be measured with various lab-based measures requiring maximum exertion. However, in the performance of a real world task, an individual might *utilise* these resources to different degrees depending on the task at hand. This resource *utilisation* requires that measurements take place while the actual task is being performed, making measurement more difficult to capture and less generalisable. It would be expected that resource *utilisation* would be highly correlated with overall task performance, but the measures that are available and commonly used are those of *resource availability*. Kondraske argues that linear combination methods (using linear regression) of measured resource *availabilities* therefore do not capture the non-linear threshold effects predicted by resource economic models, and it is the reason for the marginal performance of linear, statistically based models.

Resource economic mathematics: Mathematically, a resource economic model depends on the comparison of a set of resource availabilities (RA) with a corresponding set of demands (RD) on those availabilities derived from the high level task in question. By comparing the availabilities to the demands ($RA \geq RD$) for each resource in the set of BEPs, and combining the results using a logical AND operation, a prediction of high-level task performance could be derived. This method checks that the demands of the high-level task fall within the user's multidimensional performance envelope and then returns a negative result if *any* particular basic resource is found to be a limiting factor. This resource economic model has the added advantage that it can be applied at different hierarchical levels in a generic modelling strategy.

Figure 4-7 shows in summary form the distinction between linear combination and resource-economic models. In the top of the diagram, a linear combination model assumes that different proportion or weights (b_1, b_2, b_3 etc.) of low-level resource availabilities (RA) combine to give the level of performance on a high level task. In the bottom of the diagram, a resource economic model assumes that the high level tasks can be broken down into a set of resource demands (RD1, RD2, RD3) that correspond to a set of resource availabilities (RA1, RD2, RD3). An estimate of high-level performance can obtained by checking that *all* the availabilities are equal to or greater than the demands of the high-level task. If one resource availability, in this case RA3, is less than the demand RD3, the AND condition is not satisfied and RA3 becomes the limiting resource.

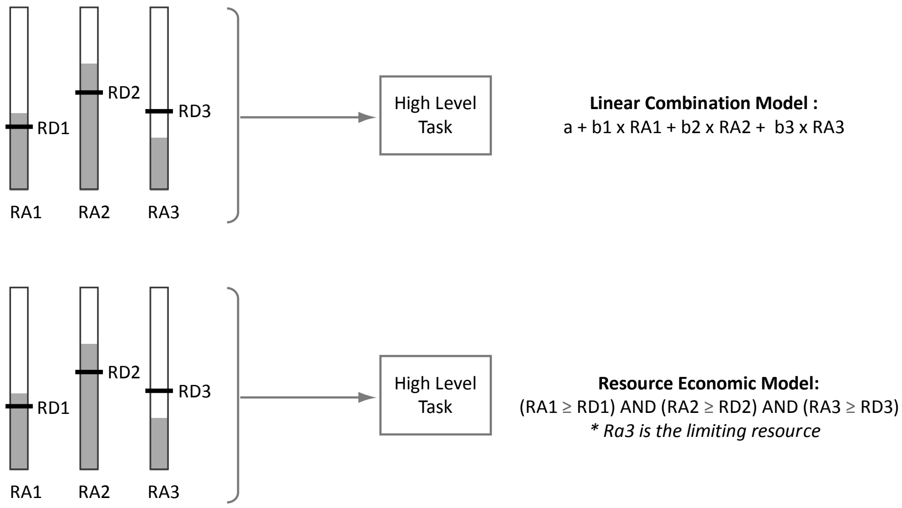


Figure 4-7 Linear combination models and resource economic model compared

A metric for performance capacity stress could be calculated which is the ratio of a resource demand (RD) to a corresponding resource availability (RA) i.e. RD/RA * 100. The threshold for a particular resource will be 100%, with values less than 100% indicating the amount of stress on a performance capacity, and values over 100% indicating the extent to which the resource demand exceeds the resource availability. Reserve capacity could also be defined as the difference between resource utilisation and resource availability i.e. RA - RD, given that RA > RD.

Thus, using the ERM involves determining the set of basic resources that can predict higher level task performance to form the basis of an evaluation method. This implies that it is more important to capture the wide range of low level resources than a few resources in depth, because this increases the probability of finding the limiting resources (Kondraske, 2006a).

Resource substitution: To explain human adaptability and coping strategies, Kondraske suggests the principle of *resource utilisation flexibility* with *resource substitution*. In this case, a performance resource may be substituted for another with the same dimensionality. Human resource subsystems will act in such a way as to minimise the stress on all resources while maximising reserve capacity. This real-time multidimensional optimisation of performance resources practically leads to the scenario where individuals with different resource profiles could accomplish a task by using different procedures.

Kondraske describes a practical implementation of these ideas via Nonlinear Causal Resource Analysis (NCRA) (Gettman et al., 2003 ; Kondraske, Johnston, Pearson, & Tarbox, 1997). In this method, a range of BEPs are measured, and users then perform a high level task of interest. Instead of looking at correlations and regression models between BEPs and overall performance in the high level task, the NCRA methodology analyses the relationship between each BEP and overall performance and fits a curve to the lower boundary of points in each plot to yield a resource demand function (RDF). This RDF therefore represents the minimum required amount of a BEP in order to achieve a certain level of high level performance. This method therefore differs from linear methods of model building by recognising the non-linear relationships inherent in the combination of BEPs as they interface with real-world tasks.

Resource variation with time: The concept of *dynamic diversity* was discussed in Chapter 2, and refers to the changing of resource capacities with time. This essential element is also built into the ERM where a person's performance envelope changes with time. This change can be brought about suddenly via disease, injury or trauma. It can also change progressively with time via the ageing process. In some cases, it can also vary on a day to day basis for conditions such as arthritis where joint pain and stiffness reduce performance resources such as strength, speed and range of motion.

4.3.4 Comparing the IDC and the ERM

The Inclusive Design Cube (IDC) was presented in Chapter 2 as a conceptual model for understanding how the range of user sensory, cognitive and motor capability in the population could be mapped to the types of product design solutions. The IDC was developed based on the Model Human Processor's sensory, cognitive and motor processors. It represents a population view and uses three broad dimensions to characterise user capability. The IDC is also disability centric with the origin of the axes representing some population average level of sensory, cognitive and motor performance. The axes point to increasing levels of *disability*, or minimal levels of performance at the extremes. The idealised volume represents the included population that will be sufficiently able to use a given design (shown on the left of Figure 4-8). In other words, users with capabilities that lie within the cube could use the design and are therefore *included*, while users with capabilities that lie outside the cube cannot use the design and are therefore *excluded*. Therefore the IDC suggests a *maximisation of the inclusion volume*.

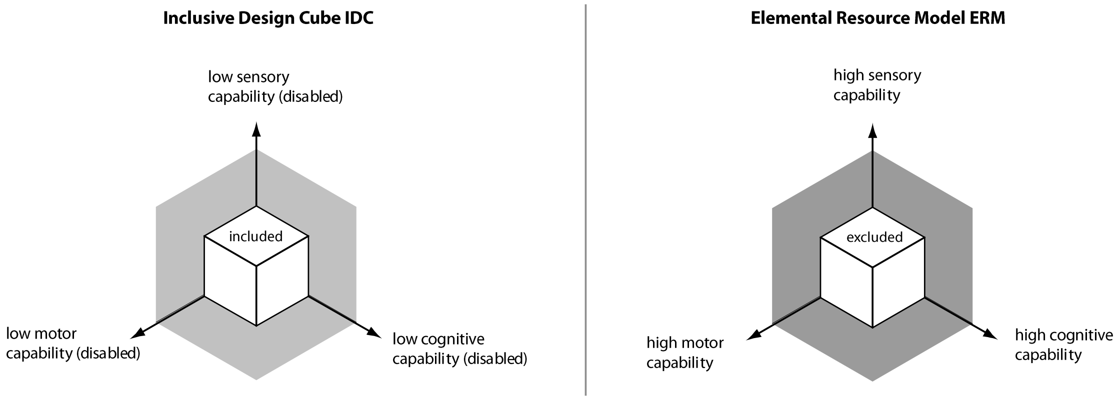


Figure 4-8 A comparison of the Inclusive Design Cube (IDC) and the Elemental Resources Model (ERM)

The previously discussed ERM is performance centric with the origin of the axes defined at zero levels of performance. The axes point to increasing levels of performance capability, or maximal levels of performance at the extremes (shown on the right of Figure 4-8). Because of the way the ERM is defined as a performance envelope, the idealised volume represents the excluded population who will not able to use a given design. Users with capabilities that lie within the volume of the cube cannot use the design and are hence *excluded*, while users with capabilities that lie outside the cube could use the design and are hence *included*. Therefore the ERM suggests a *minimisation of the exclusion volume*.

In comparing the IDC to the ERM, it is evident that they are *inverses* of each other. The IDC is also a performance envelope - just inversely defined. The boundary surfaces of both cubes are fuzzy, as in actuality there is no binary cut off point as the diagrams may suggest. The cube boundaries are the most interesting from a design viewpoint, as the aim of evaluation and redesign would be to shift these boundaries to either maximise the inclusion volume (IDC) or minimise the exclusion volume (ERM). It is expected that at these boundaries, users would experience high levels of difficulty and frustration. The relationship between a user's position in the performance space and their levels of experienced difficulty is not known specifically, as it may vary from person to person depending on their capability profile (and other factors) at the time of use.

The IDC and ERM differ in the underlying model for subsystem integration for predicting high level task outcomes. Since the development of the IDC is linked to disability scale development, it assumes a linear combination model (Martin & Elliott, 1992). The ERM utilises a non-linear combination model based on resource economic mathematics for making its predictions.

4.3.5 Section Summary

In summary, the preceding discussion demonstrated the two dominant user performance models upon which analytical evaluation methods could be based. There are traditional linear models of capability combination, and then there are non-linear models of the interaction between elemental resources of human performance capacities. The essential question remains as to which of these two underlying models more accurately represents disability phenomena and ultimately which should be adopted as the underlying theory of Inclusive Design evaluation?

Linear models based on correlation and linear regressions are well understood, and are relatively straightforward to apply compared to non-linear methods. However, their predictive capabilities, especially for non-homogenous populations, have been questioned (Kondraske, 2006a; Kondraske, 2006c). Conversely, the ERM and its associated methodology of NCRA is a relatively new development, and further testing, development and comparison is required to demonstrate its advantages. The use of the ERM also implies the measurement of a large number of elemental resources, which increases the probability of finding limiting resources for given tasks. From a practical standpoint, this may limit efforts to use it as a base for evaluation as it would require significant time and resources to collect representative data sets, and also it would require computational support for application (Kondraske, 2006b; Kondraske, 2006c). Even so, the ERM may prove useful even if it operates on a reduced set of elemental resources that are found to be responsible for limiting success in the majority of tasks with consumer products.

4.4 A Generic Analytical Evaluation Framework

Given the analytical evaluation methods outlined and reviewed in Chapter 2, this section presents a generic high-level analytical evaluation framework for Inclusive Design. The aim is to provide the foundation for a systematic, quantitative and predictive framework that would allow designers and decision makers to ask and answer questions about the inclusivity of their product designs in the design process. Based on an examination of the advantages and limitations of current methods, and studies into the needs of designers, a list of six requirements with corresponding research activities was generated as shown in Table 4-1.

Table 4-1 Requirements for an analytical evaluation framework

Requirements	Research Activities
1. General application to consumer products including products ranging from household electronic appliances to communication and information devices.	Classification of consumer product interfaces.
2. Comprehensive model of user capabilities encompassing the most common tasks that users perform with products.	Classification of human capabilities required for interaction and models of their interaction.
3. Facilitates valid estimation of exclusion and the numbers of people who are likely to experience various levels of difficulty under explicit and justified assumptions.	Establishment of an integrated user capability database with an accurate model to calculate/predict task outcomes.
4. Sensitivity to variations in product attributes that can map to changes in excluded population.	Establishment of the required precision for sensitivity analysis.
5. Usable by designers and other stakeholders, allowing ease of calculation of exclusion, visualisation to support decision making, and flexibility in using the framework at various points in the design process.	Implementation of tools incorporating calculations and different forms of visual output for different stakeholders.
6. Accommodates revisions, upgrades, extensions and new data based on usage and feedback.	Refinement and expansion based on validation studies and new research.

The requirements deal with setting out the scope, underlying model, predictive ability, sensitivity, usability and adaptability of a generic framework for Inclusive Design evaluation. These general parameters gave rise to a structured iterative evaluation process with three stages, as shown in Figure 4-9.

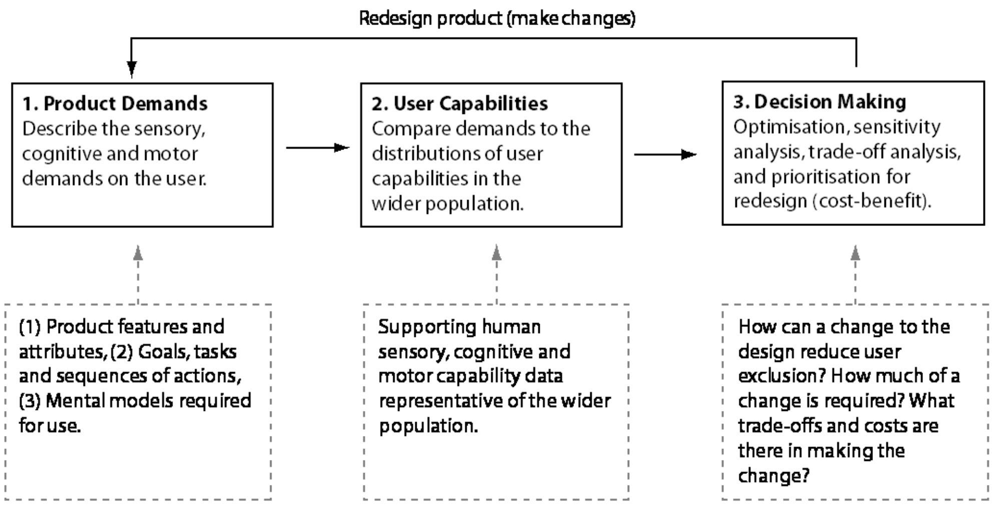


Figure 4-9 A Generic Analytical Evaluation Framework

The framework shows three stages of (1) establishing product demands on user capabilities, (2) comparing these demands to the distribution of capabilities in a population of interest and

(3) using decision making techniques to improve the design of the product. Once changes are made to the design of the product, the profile of product demands changes, and thus the process is repeated. The first stage would involve the description and representation of three components: (1) the product interface features and attributes, (2) user goals, tasks, and sequences of actions and (3) a representation of the mental models required for using the product. These three components comprise the functional demands that the product places on the user's sensory, cognitive and motor capabilities in a given use environment.

The second stage would involve the estimation of the proportions of people in a target population that may be excluded or have difficulty with the product design. This is achieved by comparing the demands to capability measures stored in a comprehensive capability database. By comparing these demands to distributions of capability levels in the wider population, an estimate of design exclusion can be obtained. Importantly, these comparisons should be multivariate and simultaneous across user capability domains. Considering the previous discussion on models of capability combination, the important factor in generating valid predictions would be the selection of a suitable model (linear, resource economic etc.) upon which to base these calculations of exclusion.

The third stage involves decision making and analysing user exclusion estimates via sensitivity and trade-off analysis. Sensitivity analysis in this case will comprise asking 'what-if?' questions about design attributes and looking at the effects of making changes to these attributes on excluded population estimates. For example, a designer might find that a significant proportion of a population of interest is being excluded by the 8 point text size on a product interface (at a given viewing distance). By increasing the text size to 10 point, the change in excluded population could be recalculated given adequate data. Thus the designer can see not only what must be done to include more people in terms of increasing text size, but also *by how much* it should be increased to achieve a given reduction in exclusion.

Trade-off considerations are also important, such as the impact of increasing text size on the aesthetics of the product or other constraints such as button size that limit the size of the text. In essence, the framework emphasises the making of informed decisions about design features and prioritises design problems based on objective user capability data. Due to the reliance on a capability database and the quantitative nature of the method, computational support will be required for a full implementation of such a framework. In addition, methods of visual decision making could be incorporated to make the framework usable. As Meister points out, the value of analytical methods could be found in the process itself, despite the final results (Meister, 1995). Re-envisioning an analytical evaluation framework for Inclusive Design as a non-prescriptive decision making tool may have advantages in terms of its acceptance and employment in industrial practice.

The critical precursor to the implementation and testing of such an evaluation framework, however, is the exploration and selection of an accurate model of capability combination which can predict task outcomes to an acceptable degree of accuracy. The studies presented in the following chapters of this dissertation aim to explore the relationships between capability measures and task outcomes in the context of requirements 1 and 2 in Table 4-1.

4.5 Chapter Summary

This chapter examined the foundations of capability-demand theory and issues with regard to modelling human functional capacity. Both linear models based on correlation/regression and non-linear resource economic models were found to be options for modelling the combination of user capabilities in relation to task performance. A three stage generic framework for analytical inclusive evaluation was also presented. The chapter ended with the need for further investigation into the nature of the relationships between user capabilities and task outcome measures. Understanding such relationships is key to the further development of the analytical framework and populating it with capability data.

Chapter 5 Capabilities and Demands

5.1 Chapter Overview

This chapter presents a review of the basic set of human capabilities utilised when interacting with everyday consumer products. The chapter begins with a description of the human sensory, cognitive and motor capabilities that account for the majority of the variation in human functioning due to ageing and disability. The review is based on literature reviews and consultations with domain experts. The demands made by products on these capabilities are also discussed. The chapter concludes with a study on the prevalence of disability and health conditions in the UK population. This is explored via secondary data analysis of the 1996/97 Great Britain Disability Follow-up Survey (DFS) dataset.

5.2 Review of Human Capabilities

Systematic reviews of functional classifications and experimental studies in the literature were undertaken in order to identify the relevant low-level capabilities relevant to product design in each of the sensory, cognitive and motor domains. This was reinforced with consultations with experts in medicine, psychology, rehabilitation and disability. The capabilities described in the following sections aim to be comprehensive in so far as they are relevant to common tasks with products (Kondraske, 1988).

Table 5-1 lists the main body of literature and design guidelines selected for review. The literature review exercise involved extracting all of the capabilities mentioned into a running list of sensory, cognitive and motor capabilities. Some literature sources aimed to be design relevant, thus focusing on the key capabilities required for product interaction. Other sources aimed to be an exhaustive classification scheme inclusive of all domains of health and disability, for example the International Classification of Functioning, Disability and Health (ICF) (World Health Organisation, 2001). The capabilities described in the following sections aim at an adequate middle ground that covers the essential capabilities that impact most product tasks.

Table 5-1 Main review papers on user capability for Inclusive Design

Author and year	Title
(Winter & Lemke, 2005)	MED-AUDIT Impairment Categories: Working towards Mapping AMI Usability
(Fisk et al., 2004)	Designing for older adults: Principles and Creative Human Factors Approaches
(World Health Organisation, 2001)	International Classification of Functioning, Disability and Health
(Jacko & Vitense, 2001)	A review and reappraisal of information technologies within a conceptual framework for individuals with disabilities
(Petrie, 2001)	Accessibility and usability requirements for ICTs for disabled and elderly people: a functional classification approach
(Hawthorn, 2000)	Possible implications of aging for interface designers
(Charness & Bosman, 1994)	Age-related changes in perceptual and psychomotor performance: Implications for engineering design.
(Morris, 1994)	User interface design for older adults
(Haigh, 1993)	The ageing process: a challenge for design
(Fisk, 1993)	Design for the elderly: a biological perspective
(Charness & Bosman, 1992)	Human factors and age
(Carmichael, 1999)	Style guide for the design of interactive television services for elderly viewers
(Story et al., 1998)	The Universal Design File - Designing for People of All Ages and Abilities
(Carroll, 1993)	Human cognitive abilities: a survey of factor-analytic studies
(Vanderheiden & Vanderheiden, 1992)	Guidelines for the Design of Consumer Products to Increase Their Accessibility to People with Disabilities - Working Draft 1.7
(Pirkl & Babic, 1988a) (Pirkl & Babic, 1988b)	Guidelines and Strategies for Designing Transgenerational Products: A Resource Manual for Industrial Design Professionals
(Fleishman & Quaintance, 1984)	Taxonomies of human performance - The description of human tasks

5.3 Sensory Functions

The sensory functions of vision, hearing, taste, touch and smell allow for the intake of information from the world. In this section, the underlying capabilities of the *distance* senses will be discussed i.e. vision and hearing. The other senses, though important for certain types of interaction, are not as critical as vision and hearing for consumer product use.

5.3.1 Vision

Various medical conditions such as age-related macular degeneration (AMD), cataracts, glaucoma, retinitis pigmentosa, and diabetic retinopathy cause reductions in the functional capabilities of the eyes making visual tasks with products more difficult (VisionConection, 2006). The reduction in the following five functions seems to account for most of visual disability (Fletcher & American Academy of Ophthalmology., 1999; Jacko, Dixon, Jr, Scott, & Pappas, 1999; Norton, Corliss, & Bailey, 2002; Schieber, 1992).

Visual Acuity: Acuity is the ability of the eyes to resolve fine details and differentiate different parts of the visual field from each other (Schiffman, 2000). There are various forms of acuity, including detection, vernier, resolution, recognition, and dynamic acuity (Schiffman, 2000). The most familiar is recognition acuity measured with a letter chart at the optician. Acuity is greatest in the central visual field and it is measured as the inverse of visual angle (a measure of the size of a target subtended on the retina). Measuring the *maximum visual acuity* thus gives the performance limit of detail perception under the measured conditions and this has implications for text at maximum contrast. However, because acuity is measured at high contrast and many real world conditions occur at lower contrast levels, recognition acuity has limited predictive value for spatial vision and form perception (Schiffman, 2000).

Contrast Sensitivity: Contrast is a measure of the difference in luminance between an object and its background. Contrast sensitivity is a measure of the minimum contrast that can be perceived at different spatial frequencies. Based on the channel model of visual form perception (Schiffman, 2000), real world visual stimuli are analysed by Fourier analysis into a range of spatial frequencies at various contrast levels. Large objects comprise low spatial frequencies and small objects comprise high spatial frequencies. An individual’s sensitivity to contrast at these different frequencies can be reduced due to ageing and various eye conditions, leading to reduced performance in form perception and mobility.

Sine wave gratings are used to assess an individual's contrast sensitivity at different spatial frequencies and orientations, and the results are plotted on a graph of contrast sensitivity versus spatial frequency. This graph is termed the contrast sensitivity function or CSF (Schiffman, 2000). An example of the CSF is shown in Figure 5-1. The figure shows the reduction in contrast sensitivity for medium and high spatial frequency stimuli in the 80-year age group compared to the 20-year age group. For older people, the loss of contrast sensitivity at high spatial frequencies is generally greater than the loss at low spatial frequencies. The shaded region of the graph shows the range of high spatial frequencies at which visual acuity is measured. This demonstrates that visual acuity measures at high spatial frequencies cannot predict performance on real world vision tasks because it does not capture contrast thresholds for larger objects of lower spatial frequencies. The CSF therefore characterises a person's *maximum contrast thresholds* for various sizes of objects, and it can be used to predict visual exclusion of product controls and other form features.

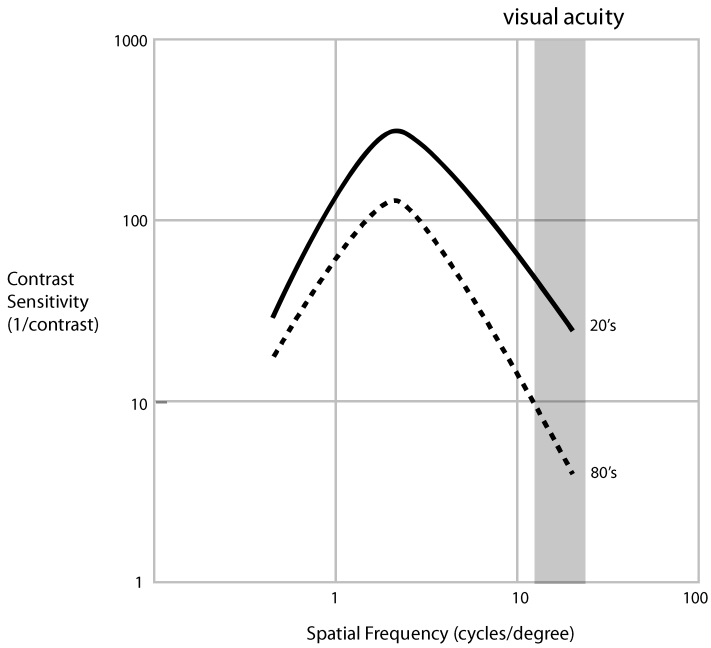


Figure 5-1 Visual acuity and contrast sensitivity with age

Colour Perception: The eyes are equipped with two types of photoreceptors known as rods and cones. In the human eye, there are three types of cones that are sensitive to short, medium and long wavelengths of light. The operation of these three types of cones gives rise to colour perception. People exhibit colour blindness if any of these cones are missing or defective, leading to a loss of discrimination of the full spectrum of colour. Despite the name, complete colour blindness, where there is no perception of colour, is extremely rare. There are two main forms of colour blindness: red-green and blue-yellow. This means that for red-green colour blindness, a person cannot distinguish colours between red and green in the colour spectrum and for blue-yellow colour blindness, a person cannot distinguish from the yellow to the blue part of the colour spectrum. Colour discrimination is also known to decrease with age. Colour blindness does not cause significant problems provided that foreground and background colours are of sufficient contrast to be detected. However, if colour is used to display information then the possible range of colour blindness and colour confusions becomes an important consideration. Therefore, the *maximum* range of colours discriminated becomes the threshold for colour perception.

Central and peripheral field of view: Various conditions can cause a reduction in the useful field of view. The central visual field can be obscured or the peripheral visual field can be reduced resulting in 'tunnel vision'. Partial combinations of central and peripheral loss are also possible. Loss of central field of view is more difficult to accommodate because acuity is greatest in the central field for colour vision. For design evaluation, the *maximum extent* of the central field of view is the important measure. The size and layout of interface features can be evaluated to determine if they fall easily within the users' field of view when performing essential tasks. By grouping controls and designing them in close proximity, the product would not demand a large field of view for search and detection of the relevant controls.

Depth perception and stereopsis: Stereopsis is the ability to perceive depth based on the combination of two slightly varying images transmitted to the brain by the spherical geometry of the eyes. Depth information is also obtained from motion through space. People with loss of depth perception can have difficulty operating in a three dimensional environment. Movement is affected and products that require spatial manoeuvres to access controls could be especially difficult.

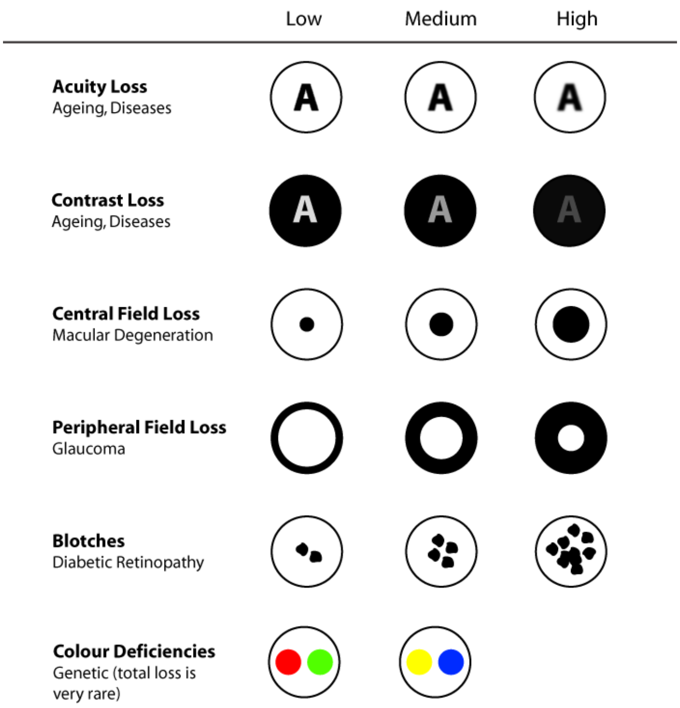


Figure 5-2 Visual capability losses

The diagram in Figure 5-2 symbolically demonstrates the effects of these losses of acuity, contrast, central field, peripheral field, available field and colour range in functional vision. The reality is that there may be simultaneous multiple loss, which results in a superposition of these losses on functional vision. Common visual tasks include detecting and reading text on the product chassis and displays, detecting symbols and graphics, and detecting features of the product against the background of the chassis. Design parameters such as size, shape, colour, and contrast all constitute visual demands on the user.

5.3.2 Hearing

Hearing capabilities are important for product interaction when using products that provide auditory output and facilitate speech communication. Loss of hearing capability can be attributed to three main mechanisms: conductive hearing loss, sensorineural hearing loss and a mixture of both types of hearing loss (Moore, 1998; Moore, 2003; Schiffman, 2000). Common conditions include presbycusis (loss of hearing with age), tinnitus (ringing in the ear), loudness recruitment (sounds become uncomfortably loud as they are increased) and loss of frequency selectivity (adequately filtering frequencies in sound spectra).

The less prevalent conductive hearing loss results in the attenuation of incoming sound in a relatively uniform manner across all sound frequencies. This is perceived as a reduction in loudness of the incoming sound stimulus. Conductive hearing loss is relatively easy to remedy and involves increasing the sound levels by the amount it was attenuated. However, sensorineural hearing loss is more prevalent in the population, and it results in variable losses in threshold in different frequencies of the human frequency range (Moore, 1998). This loss can occur at any of the frequencies and results in problems with sound discrimination.

Hearing capability is typically characterised by pure tone audiometric testing resulting in an audiogram. Figure 5-3 shows an audiogram with plots of a conductive hearing loss profile and a sensorineural hearing loss profile. The conductive profile is relatively flat with equal loss in threshold across all frequencies whereas the sensorineural profile shows large losses in threshold for higher frequencies. Presbycusis is an age related condition that results in the loss of higher frequencies, affecting the discrimination of sound in noise. Another phenomenon of importance is loudness recruitment, where the rate of increase in loudness level is dependent on overall loudness and may be greater in impairment (Moore, 1998).

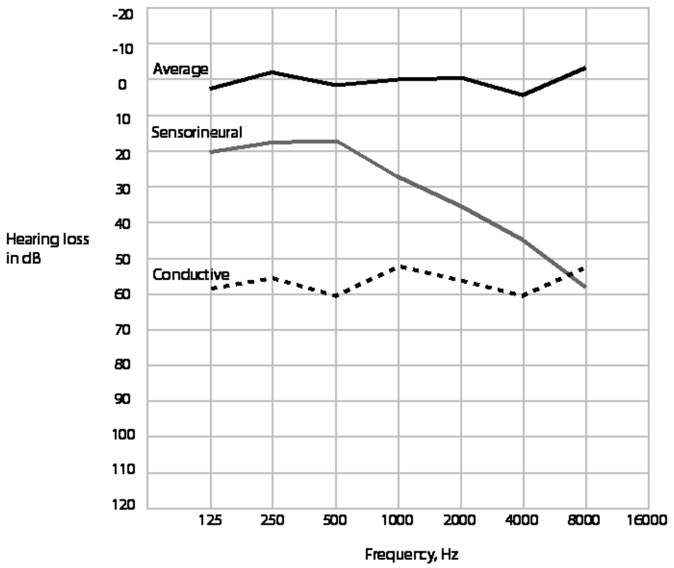


Figure 5-3 An example of an audiogram

The following three low-level hearing functions are necessary for product interaction:

Pure Tone Detection Thresholds: For determining sound detection, audiogram measures of intensity thresholds for different frequencies can be used to predict whether a sound will be detected (DeBonis & Donohue, 2004). Auditory output is either constrained to tones and beeps at a particular frequency or the output has a complex spectrum such as speech and music. By comparing the audiogram thresholds at various frequencies to the frequency spectrum of the sound output, an estimate of exclusion for simple tones and areas of difficulty for complex sounds (such as the effect of high frequency loss on speech perception) could be obtained.

Speech detection and recognition discrimination thresholds: The speech detection threshold (SDT) and speech recognition thresholds (SRT) are useful measures in addition to pure tone thresholds to characterise hearing loss. These give measures of a person's ability to detect and understand speech. There is a difference between the SDT and the SRT i.e. the speech recognition threshold is usually higher than the speech detection threshold (DeBonis & Donohue, 2004).

Sound localisation: Localisation is the ability to tell the direction from which sound is coming, and this ability reduces with age. It is based on the timed phase difference of sounds entering each ear. It is an important consideration when designing products that signal the user such as alarms, mobile phones and public address systems.

Most hearing tasks with products involve detecting, discriminating and localizing tones, speech and music. Design parameters include the frequency spectra of the sound, the overall loudness and the extent of competing background sounds. Vision and hearing capabilities are similar in that the measure of visual contrast is similar to the measure of loudness. Visual and auditory signals can both be characterised by spatial frequency profiles, as in the contrast sensitivity function (CSF) and sound frequency profiles of the audiogram, respectively. These capability profiles are the main predictors of real world signal detection tasks.

5.3.3 Environmental Factors and Sensory Capabilities

Environmental factors can affect the detection of visual and auditory stimuli. Visual capabilities can be affected by ambient illumination which can cause glare problems. Surface finishes that are highly reflective can also easily cause glare reflections. The level of illumination determines the luminance values of foreground and background features on the product, leading to reduced contrast of product features. Hearing capabilities can be affected by noise (unwanted sound) which can cause problems in speech detection and discrimination - especially for older people. Noise can also mask various frequencies in the target stimulus. For example, hearing speech in a background of speech noise can be particularly difficult when performing listening tasks in a restaurant.

Measurement of vision and hearing capabilities are usually obtained in a standard testing environment that does not reflect real world conditions. Normal assumptions about the operating environment are necessary when trying to relate product demands to capability measures. If the actual operating environment is close to the conditions under which sensory capability is measured, the error in using such measures can be assumed to be minimal. For products that operate in a range of environments, such as mobile phones, worse case assumptions about levels of illumination and noise can be used to relate the product demands to capability measures. Therefore additional measures of visual capability in low light environments and hearing capabilities in noisy environments will be required.

5.4 Cognitive Functions

Various cognitive architectures have been proposed in the literature that attempt to describe the information processing sub-systems involved in cognition, such as ACT-R, SOAR, and EPIC (Adams, Langdon, & Clarkson, 2002; Byrne, 2003; Proctor & Vu, 2003; Wickens & Hollands, 1999). Mainstream cognitive psychology attempts to generalise the architecture of the mind for broad application, and a distinction can be drawn between the *architecture* of the cognitive system and the *contents* of the cognitive system at different points in time (Payne, 2003). Wickens describes a model of human information processing stages shown in Figure 5-4 (Wickens & Hollands, 1999).

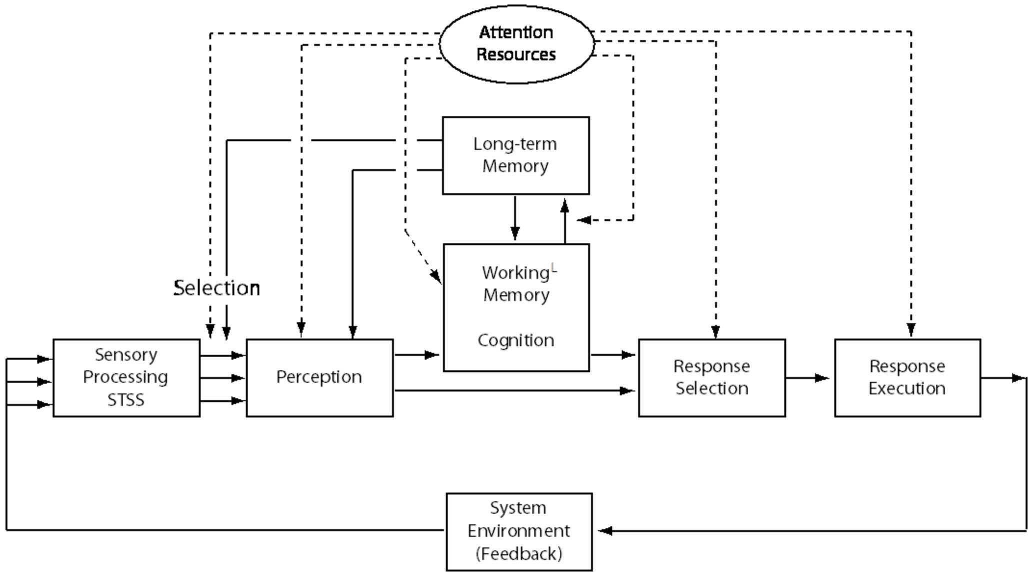


Figure 5-4 A model of human information processing stages (Wickens & Hollands, 1999)

The model is a stage based model i.e. it consists of a series of operations on information. In addition, the processing may begin at any stage by external input or intentions to act (Wickens & Hollands, 1999). The various stages of the model will be discussed in the following sections.

5.4.1 Sensation and Perception

Starting at the left of the diagram, sensory information is transmitted to the brain via initial sensory processing. There is a short-term sensory store (STSS) associated with this stage that stores the sensory input for a short time. The quality of information reaching the brain depends on the state of the sensory receptors. As previously discussed, if there are reductions in visual and auditory capabilities, the raw sensory information reaching the brain is degraded, and this impacts on the stages that follow.

Raw sensory input is decoded and interpreted in the next stage of perception. Perception is both bottom-up and top-down in that characteristics of the input signal together with expectations from long term memory are used to rapidly categorise the incoming information. The process is rapid and requires less mental effort and processing time than cognitive operations using working memory. The bottom-up and top-down perceptual processing is adaptive, so that if the incoming sensory information is poor for example, more weight will be given to the top-down processing utilising stored experiences in long term memory.

5.4.2 Working Memory

Working memory can be considered as a temporary store for activated information (Baddeley, 2000). Baddeley proposes that it is structurally organised by different modalities of storage (Baddeley, 2000; Baddeley, 2002). This working memory model is shown in Figure 5-5. It consists of three subsystems: the visuospatial sketchpad, the episodic buffer and the phonological loop. The central executive is responsible for attentional control and coordination of the working memory subsystems. The working memory subsystems are linked to long term memory where learned information is retrieved. The working memory system is responsible for cognitive tasks such as rehearsal, reasoning, image transformation, planning and problem solving (Wickens & Hollands, 1999).

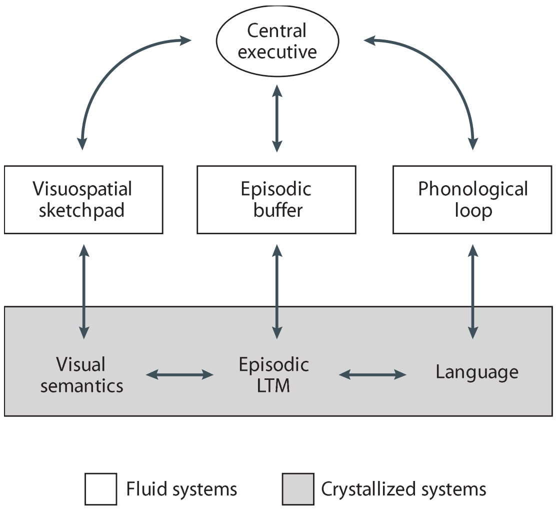


Figure 5-5 A model of working memory

Working memory has been found to have a limited capacity for stored information with a duration of 10-15 seconds. The general capacity of the working memory system has been estimated to be around five to nine chunks of information (Baddeley, 2000). However, more complex items such as sentences, procedures, or images can be remembered as if they were individual elements when chunked. This is after prolonged use has caused them to become well established in the more permanent store of long term memory. Another important characteristic of working memory is that the central executive is assumed to have limited resources of attention. This can be overloaded by either increasing the volume of individual items to deal with or the number of simultaneous activities that require attention. Therefore working memory and attentional capacities are the limiting factors when interacting with products. Thus two important performance measures for the working memory system are *storage capacity* in terms of the number of chunks that can be held and *speed and accuracy* of processing.

5.4.3 Long Term Memory

Long term memory is a permanent store for knowledge gathered from experience. The type of knowledge stored can be classified into various types including semantic memory, episodic memory and procedural memory. It is useful to distinguish between knowledge of product features and how the product *works*, versus knowledge on how to *use* the product in terms of action sequences that will achieve goals. These two types of knowledge are inter-related, and are both used when interacting with products. Therefore, recognition and recall capabilities are the limiting factors in the performance of long term memory. Measures of recognition capability are required when comparing visible product features to information stored in long term memory. Recall capability measures are needed for determining users' ability to retrieve stored knowledge about product features and behaviour.

5.4.4 Mental Models

Users are assumed to construct mental models in working memory based on previous knowledge cued by current environmental characteristics and use this representation as they proceed through the interaction (Cañas, Antolí, & Quesada, 2001; Van der Veer & Melguizo, 2003). These models may reflect their understanding of the behaviour of the product and how it is to be used (Norman, 1983; Van der Veer & Melguizo, 2003). The concept of mental models has received significant attention in the HCI and applied cognitive science literature. It has been found that mental models can be incomplete, unstable, unscientific, and parsimonious while varying in complexity depending on the degree of previous experience (Norman, 1983). Because of this dependence on previous experience and continuous modification through successive interactions, mental models can be difficult to capture (Van der Veer & Melguizo, 2003).

From a practical standpoint, the concept of a mental model could be used to enable the operational estimation of cognitive information processing demands. Mental models could be embodied in a *representation* that captures a *demanded* mental model of device usage. Cognitive processes therefore could be assumed to act on this representation during the guidance of action. A mental model representation consists of knowledge of the various interface features and how they work, and a representation of the action sequences necessary for moving from an initial state to a goal state. Because working memory is limited, there will be a limit to the complexity of the mental model in mind and the mental operations that can be performed on it at any one time. In summary, interaction involves retrieving previous knowledge of a particular product's features and how the product works, and using this knowledge together with task demands and product perception to form and operate on mental models in working memory (Van der Veer & Melguizo, 2003).

5.4.5 Response Selection and Execution

Cognitive processing results in the selection of a response that would move a user toward a goal. Response selection is distinct from response execution, as execution involves the coordination of movement. Even though a given response selection may be correct, the execution may be erroneous (Wickens & Hollands, 1999). Successful execution also depends on the state of the effectors i.e. the motor capabilities of the user.

5.4.6 Language and Communication

Language and communication capabilities involve the comprehension and expression of verbal and written language. An assessment of language comprehension capability is necessary when performing tasks such as reading labels and product manuals. It is also employed to interpret verbal messages from a product or system. An assessment of language communication capabilities is important for giving spoken commands to a product. For product evaluation, the primary concern is with linguistic communication in speech and sentence construction as these are most commonly employed in product design. Language and communication capabilities depend on the perceptual, working memory and long term memory systems (Wickens & Hollands, 1999).

5.4.7 Age Related Effects

The literature shows that ageing is in general accompanied by declines in the cognitive capabilities mentioned above including attention, memory, reasoning and problem solving. However, certain forms of crystallised intelligence, for example vocabulary, do not show decline with age (Park, 1999). The effects of conditions such as dementia, Alzheimer's, brain injury, cerebral palsy, and amnesia are such that cognitive decline is either accelerated, or certain cognitive capabilities are lost. The aforementioned cognitive capabilities form the infrastructure on which cognitive task performance is based, and therefore should constitute the set of limiting factors to successful product interaction.

5.5 Motor Functions

As previously mentioned, response execution involves taking action in the world by the controlled movement of the effectors of the human body. Product interaction requires operating product controls and moving around in the usage environment. Physical function can be divided into six main areas: muscle performance, cardiopulmonary/endurance, mobility/flexibility, neuromuscular control/coordination, stability and balance/postural equilibrium (Kisner & Colby, 2007). Definitions for these functions are given in Table 5-2.

Table 5-2 Definitions of the aspects of physical function (Kisner & Colby, 2007)

Aspect of Physical Function	Definition
1. Muscle performance	The capacity of muscle to produce tension and do physical work. Muscle performance encompasses strength, power, and muscular endurance.
2. Balance/Postural Equilibrium	The ability to align body segments against gravity to maintain or move the body (centre of mass) within the available base of support without falling; the ability to move the body in equilibrium with gravity via interaction of the sensory and motor systems. Postural control, postural stability, and equilibrium are used interchangeably with static or dynamic balance.
3. Cardiopulmonary/Endurance	The ability to perform low-intensity, repetitive, total body movements (walking, jogging, cycling, swimming) over an extended period of time.
4. Mobility/Flexibility	The ability to move freely, without restriction. The ability of structures or segments of the body to move or be moved in order to allow the occurrence of range of motion (ROM) for functional activities (functional ROM). Passive mobility is dependent on soft tissue (contractile and noncontractile) extensibility; in addition, active mobility requires neuromuscular activation.
5. Neuromuscular control/ Coordination	Interaction of the sensory and motor systems that enables synergists, agonists and antagonists, as well as stabilizers and neutralizers to anticipate ore respond to proprioceptive and kinesthetic information and, subsequently, to work in correct sequence to create coordinated movement. The correct timing and sequencing of muscle firing combined with the appropriate intensity of muscular contraction leading to the effective initiation, guiding, and grading of movement. It is the basis of smooth, accurate, efficient movement and occurs at a conscious or automatic level.
6. Stability	The ability of the neuromuscular system through synergistic muscle actions to hold a proximal or distal body segment in a stationary position or to control a stable base during superimposed movement. Joint stability is the maintenance of proper alignment of bony partners of a joint by means of passive and dynamic components.

Other similar categorisations are given in the literature, for example the five areas of posture, mobility, coordination, strength and effort, and energy utilised in moving and interacting with tasks, objects and environments (Crepeau, Cohn, & Schell, 2003). Definitions for these five functions with example actions are given in Table 5-3. These functions all relate to each other and can be measured through variables such as control precision, multi-limb coordination, response orientation, rate control, arm-hand steadiness, manual dexterity, finger dexterity, wrist-finger speed, static strength, explosive strength, dynamic strength, extent flexibility, dynamic flexibility, gross body coordination and equilibrium, and stamina among others (Crepeau et al., 2003; Jacko & Vitense, 2001).

Table 5-3 Motor Skills (Crepeau et al., 2003)

Motor Skills	Description	Example Actions
1. Posture	Relates to the stabilising and aligning of one's body while moving in relation to task objects with which one must deal.	Stabilises, aligns, positions
2. Mobility	Relates to moving the entire body or a body part in space as necessary when interacting with task objects.	Walks, reaches, bends
3. Coordination	Relates to using more than one body part to interact with task objects in a manner that supports task performance.	Coordinates, manipulates, flows
4. Strength and Effort	Pertains to skills that require generation of muscle force appropriate for effective interaction with task objects.	Moves, transports, lifts, calibrates, grips
5. Energy	Refers to sustained effort over the course of task performance.	Endures, paces

The following sections highlight pertinent issues with motor skills and its measurement for characterising motor capability loss.

5.5.1 Hand and Arm Functions

Hand function is a combination of motion control, grasping and force exertion. Most products require the use of the hands and arms to operate controls and manipulate various product features. Product interaction may demand reaching and grasping, and the exertion of linear and rotational forces with each hand separately or in combination.

Static and functional anthropometry: Anthropometry (dimensions and ranges of motion) capture the ranges of motion of the hands and arms required (distances and angles) to access a product. The capability to access and grasp product controls depends on the maximum vertical and horizontal distances that can be reached with each arm. Reaching capabilities are more important for fixed products such as washing machines and ATMs. The maximum reach distances can be measured by determining the reach envelopes in both the vertical and horizontal directions.

Grasping and force exertion: Forces can be exerted on product controls without any gripping required using non-prehensile movements (Napier, 1956). Linear forces can be exerted with fingers or the palm of the hand, for example in pushing a button. Controls and handles can also be grasped and held between fingers or within the compass of the hand using prehensile movements (Napier, 1956). Gripping actions can be sub-divided into precision gripping and power gripping. Precision grips use opposing forces of the fingers and thumb digits of the hand for fine linear or rotational movements. Power grips use the palm of the hand, in addition to the fingers, to exert larger grip forces on the product chassis and handles (Napier, 1956). Both grip types are endpoints of a continuum of grasps, and represent the movement from large force exertion with gross motions of the hand and arm to smaller force exertion with fine finger motions. Different functional grips are catalogued in taxonomies of functional grasps available in the literature (Cutkosky & Wright, 1986; Edwards, Buckland, & McCoy-Powlen, 2002; Kroemer, 1986; Napier, 1956).

Forces exerted can be divided into linear forces and rotational forces. Linear forces can be characterised in a coordinate system of three principal directions using the body as a reference point: vertically (up-down), horizontally (left-right) and ventrally (forward-back). Rotational forces can be described in both the clockwise and anti-clockwise directions. These forces need to be applied for different durations depending on the task at hand. Forces also need to be considered along with the extent of available motion within the full range of articulation. A button push requires a linear finger force exertion for a short period of time as opposed to lifting a kettle to pour which requires lifting a load for a longer time period. Measurement of the maximum linear, rotational and grip forces that can be exerted with different grasps are required to determine the maximum performance capabilities for each hand.

Dexterity: Manual dexterity is the ability to make skilful coordinated movements of one hand, a hand together with its arm, or two hands to grasp, place, move or assemble objects (Crepeau et al., 2003). Finger dexterity is the ability to make skilful, coordinated movements of the fingers of one or both hands to grasp, place or move small objects (Crepeau et al., 2003). Dexterity can be measured with tests such as the Purdue Pegboard Test (Crepeau et al., 2003), and it can be used to assess the ability to perform actions in activities of daily living (for example buttoning a shirt) or making fine manipulations with products (Tsai & Lee, 2009).

Two handed actions and coordination: Some motor actions with products are two handed operations requiring the use of both hands with different grips. Combinations of power and precision grips with force exertions are required to accomplish tasks such as opening a jar and dialling a land-line phone while holding the receiver. Two handed actions require coordination and adequate function in both hands and arms in order to successfully complete the task. The issue of handedness is also an important consideration. A product that requires the use of only the right hand would exclude or raise difficulty for left handed users.

In essence, the demanded hand actions are determined by the design of the product. There may also be multiple ways of performing an action by configuring the hands and motions. Coping strategies are important in motor actions as more freedom and ways to manipulate a product make it easier for users to adopt different hand positions to exert forces in different directions (Kanis, 1993; Yoxall et al., 2010c). Data is also available on the range of motions, forces and grips that older and disabled users can apply (Kanis, 1993; Smith et al., 2000; Steenbekkers & VanBeijsterveldt, 1998; Yoxall et al., 2006; Yoxall, Kamat, Langley, & Rowson, 2010a; Yoxall et al., 2010b)

5.5.2 Gross Body Functions

Mobility functions are necessary for consumer products such as vacuum cleaners where a user is required to move while exerting forces on the product. This includes *maximum* bending ranges that indicate the extents of upper body flexion (Crepeau et al., 2003). For products such as washing machines that require reaching into the drum, reaching and bending capabilities are also linked. Locomotion and balance capabilities are also required for walking and moving around. Performance measures such as maximum walking speed and distance can give an indication of locomotion ability. In impaired populations, various mobility aids compensate for the loss of locomotion capability. Measures of locomotion capability should include various aids if they are used on a regular basis. The type of aid used could also be captured. Wheelchairs affect the reaching and bending envelopes of users, and measures for this sub-population would also be necessary.

5.5.3 Speech Functions

The ability to produce speech depends on motor control of the vocal systems. Certain products require input in the form of voice commands, making the ability to produce coherent and clear speech a requirement (Crepeau et al., 2003).

5.5.4 Common Conditions Causing Loss of Motor Function

Conditions such as arthritis, stroke, multiple sclerosis, cerebral palsy and missing or damaged limbs can cause reductions in grasp forces, ranges of motion and fatigue thresholds in addition to the decline in these capabilities with ageing (Jacko & Vitense, 2001; Sears & Young, 2003). Design guidelines for reduced motor ability have included minimising the forces required to operate controls; not requiring simultaneous manipulations; and allowing freedom and flexibility for control manipulation, such as different grips where possible (Kanis, 1993). All products require physical interaction of some form, therefore the range of movements and exertable forces with different body effectors constitute the motor demand on users with reduced motor capability.

5.6 Capabilities and Health: A UK Population Perspective

In order to better understand the distribution of capability limitations and the prevalence of health conditions in the UK population, a secondary data analysis of the 1997/98 Disability Follow-Up Survey (DFS) data was conducted. Though population representative capability data quickly becomes outdated due to changing demographics (United Nations, 2009; Yoxall et al., 2006), this survey data was chosen for analysis because it was readily available and the measures used enabled detailed analysis of the prevalence and structure of disability in the UK population.

The 1996/97 Great Britain Disability Follow-up Survey (DFS) to the 1996/97 Family Resources Survey (FRS) used a measure of severity of disability and severity scales developed for the 1985 OPCS Surveys of Disability in Great Britain (Grundy et al., 1999). Scale items based on the ICIDH (International Classification of Impairments Disabilities and Handicaps) definitions of disability were established in the survey through the consensus of judges assessing the capability limitations of a variety of disabilities and their combinations. It should be noted that the ICIDH has now been updated to the ICF - International Classification of Functioning, Disability and Health (World Health Organisation, 2001). Table 5-4 shows the 13 disability scales together with the variable codes used in the survey. To give a measure of disability in each of the 13 categories, a one-dimensional interval scale was constructed from judgements and used as an estimator. The developed 13 scales assessed the prevalence of multiple capability loss by using a linear-additive model of influence on overall severity. The resulting alignment of the severity scales allows for the combination of scores to give an overall severity score. This was derived as a weighted sum where:

$$\textbf{disability score} = \textbf{worst} + \textbf{0.4} \times \textbf{(2nd worst)} + \textbf{0.3} \times \textbf{(3rd worst)}$$

The disability interview also collected up to four health complaints from each participant in the survey. This was recorded as textual data and each complaint was coded into 42 categories based on the World Health Organization’s International Classification of Diseases (ICD) (World Health Organisation, 1992). These categories were further grouped into 15 disease categories from the 42-way classification, and these top-level classifications are shown in Table 5-4. The 13 disability scores and 15 disease variables were obvious candidates for analysing the prevalence of capability loss and health conditions in different age groups in the UK population.

Table 5-4 Severity variables and disease variables

13 Disability Scale Variables	15 Disease Variables	42 Way Classification
Locomotion *(loc)* Reaching and Stretching *(reach)* Dexterity *(dex)* Seeing *(see)* Hearing *(hear)* Independence/Personal Care *(ind)* Continence *(cont)* Communication *(comm)* Behaviour *(beh)* Intellectual Functioning *(int)* Consciousness *(fits)* Eating, drinking and digesting *(dig)* Disfigurement *(scar)*	Infectious and parasitic Neoplasms Endocrine and Metabolic Blood and blood forming organs Mental Nervous System Eye complaints Ear Complaints Circulatory System Respiratory System Digestive System Geinto-urinary System Skin disease or disorders Musculo-skeletal System Other and Vague	Infectious Disease, Cancer, Diabetes, Other Endocrine, Blood Problem, Mental, Other Mental, Epilepsy, Migraine, Other Nervous Disorder, Stroke, Cataract, Other Eye Complaint Deafness, Tinnitus, Meniere's Disease, Other Ear Complaint, Heart Attack, Blood Pressure, Other Heart Ailment, Piles, Varicose Veins, Circulatory Disorder, Bronchitis, Asthma, Hayfever, Other Respiratory Ailment, Stomach Ulcer, Other Digestive Illness, Bowel Complaint, Teeth, Kidney Disease, Urinary Illness, Other Bladder Complaint, Reproductive Illness, Skin Ailment, Arthritis, Back Problem, Other Bone Problem, Other, Unclassified.

The sample for the DFS consisted of 7,263 participants aged 16 years and over, selected on the basis of responses to questions asked in the FRS. Missing value analysis (MVA) was first performed on the data set prior to the data analysis. Based on the results of the MVA, cases with missing values in the 13 capability and 15 health variables of interest were removed. The data set was thus reduced to 7,168 cases. The 13 disability scale variables were recoded into separate nominal variables which indicated the presence or absence of a limitation. Since the objective of this study was to explore the capability variation in the adult disabled population, cases were selected based on at least one non-zero score on any of the thirteen disability scales (which is the same as an overall severity score of 1 or more). This resulted in the final data set containing 5,704 cases. Those cases that were excluded contained data on respondents who reported one or more health condition, but no capability limitation. Every case in the data set was weighted to give population estimates for disability prevalence (Semmence et al., 1998).

5.6.1 Disability Prevalence

The proportion of people in the UK population with capability limitations and health conditions were portioned into 5 age groups are shown in Figure 5-6 and Figure 5-7. Since the survey consisted of respondents in the adult population (16 years and over), the 16-20 year group differs in width to the other age bands. In addition to demonstrating the difference in proportions of people with capability limitations and health conditions in the UK disabled population, the graphs also show differences with age.

For capability limitations, the 16-20yr group shows the highest proportions of cases for fits, communication, and intellectual functioning. The 21-40yr group also shows the highest prevalence with respect to behavioural limitations. The 81-100yr group shows the highest proportions for locomotion, reach and stretch, seeing, hearing and continence capability loss. Thus the graph shows capability limitations dealing with mental functions being more prevalent in the younger age groups (16-20 and 21-40yrs) as opposed to sensory and motor capability limitations being more prevalent in the older age groups (61-80 and 81-100 yrs). The capability proportion line for the 61-80 year group follows the 81-100yr group closely – except for seeing and hearing capability losses where the 81-100yr group shows significantly higher proportions.

Figure 5-7 shows relatively high proportions of cases of infectious, mental, skin and nervous diseases in the 40yrs and younger age groups. The 81-100yr group shows relatively high proportions of cases with blood, eye, ear, genito-urinary and other diseases. The 61-80yr age group shows the highest proportions in the neoplasm, endocrine/metabolic, circulatory and digestive disease categories. Musculo-skeletal diseases accounted for the largest proportion of people in the 41-60yr group. This was followed by the 61-80yr group and the 81-100yr group. It is evident therefore that the prevalence of different types of capability limitations and health conditions are different for different age groups.

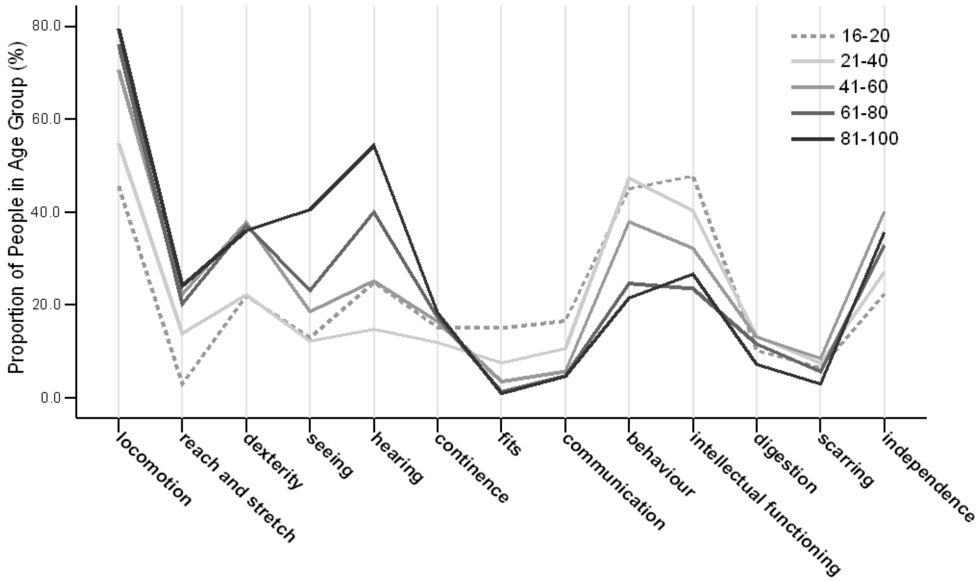


Figure 5-6 Proportions of people with capability limitations in the GB population

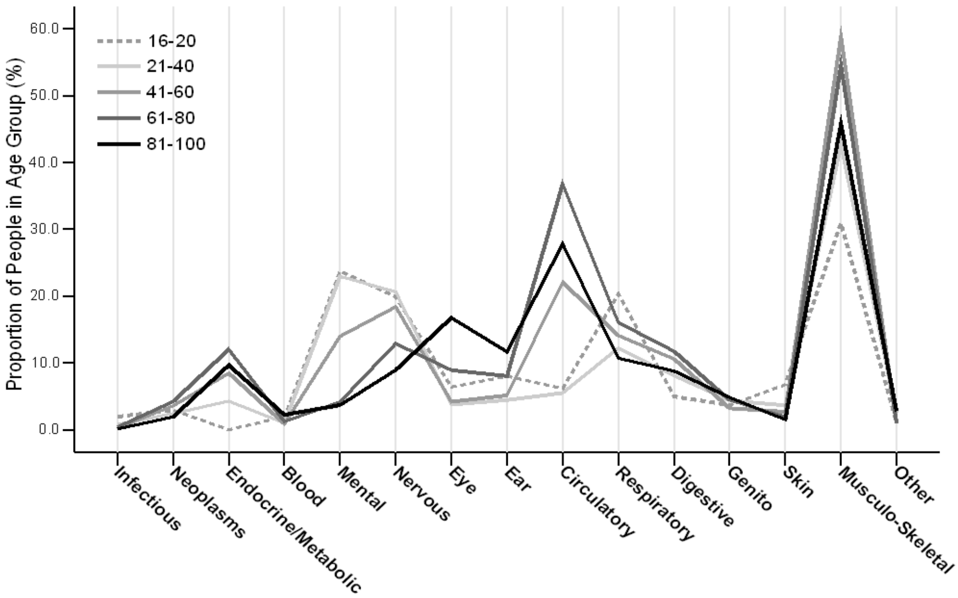


Figure 5-7 Proportions of people with a health condition in the GB population

The disabled population aged 40yrs and younger shows higher prevalence of mental and communication problems, while the 61yrs and older population shows higher prevalence of physical and sensory capability limitations compared to the younger age groups. In addition, the prevalence of musculo-skeletal diseases and locomotion problems are high for all age groups. Differential mortality between the ages of 60 and 100 could account for anomalous reversals of trend in, for example, muscular-skeletal disorders.

The high prevalence of mental and nervous diseases explains the mental capability losses in the younger age groups. In addition, the high prevalence of musculo-skeletal diseases in the 61yrs and older groups explain locomotion, reach and stretch and dexterity capability losses. The eye and ear diseases in this group explain the seeing and hearing capability losses, possibly in addition to natural degradation due to ageing. Finally the high prevalence of genito-urinary diseases can explain the high prevalence of continence problems compared to other groups. Broad relationships could thus be observed between the medical condition categories and the categories of capability loss by examining prevalence.

Table 5-5 shows the three most prevalent capability losses among seven of the disability categories that are critical for product interaction: locomotion, reach and stretch, dexterity, seeing, hearing, communication and intellectual function. These are shown together with the four most prevalent health conditions in each group. The last row of the table ranks the most prevalent capability losses and health conditions for all adults 16 years and over.

Table 5-5 Disability variables and disease variables

Age	3 most prevalent ability lossess	4 most prevalent health conditions
16-20	(1) Intellectual functioning, (2) locomotion, (3) hearing	Musculo-skeletal, Mental, Respiratory and Nervous
21-40	(1) Locomotion, (2) intellectual functioning, (3) dexterity	Musculo-Skeletal, Mental, Nervous and Respiratory
41-60	(1) Locomotion, (2) dexterity, (3) intellectual functioning	Musculo-Skeletal, Circulatory, Nervous, Respiratory
61-80	(1) Locomotion, (2) hearing, (3) dexterity	Musculo-Skeletal, Circulatory, Respiratory and Nervous
81-100	(1) Locomotion, (2) hearing, (3) seeing	Musculo-Skeletal, Circulatory, Eye and Ear
All ages (16+)	Locomotion, dexterity, hearing, intellectual functioning, seeing, reach and stretch, communication	Musculo-Skeletal, Circulatory, Nervous Respiratory, Digestive, Mental, Endocrine and Metabolic, Eye, Ear, Genito-Urinary, Neoplasms, Skin, Other, Blood, and Infectious

The prevalence of motor ability loss such as locomotion and dexterity and the high prevalence of musculo-skeletal conditions are evident across all ages. It also shows that cognitive ability loss is more prevalent in the younger age groups of 16-20 and 21-40 with associated high prevalence in mental conditions. The older age groups of 61-80 and 81-100 show high prevalence of sensory ability loss (hearing and seeing) likely due to ageing and the increased prevalence of eye and ear conditions. The data structure was further analysed by utilising hierarchical agglomerative clustering methods to explore the groupings of ability loss and health conditions.

5.6.2 Cluster Analysis

Cluster analysis is an exploratory multivariate statistical technique that is traditionally used to group and classify objects (Everitt et al., 2001). By using numerical methods of classification, it is possible to extract the underlying structure in data without any prior assumptions. Hence clustering methods are primarily descriptive and exploratory in nature (Everitt et al., 2001). It was decided that the structure of the data set could be better explored and visualised through cluster analysis, as such an investigation had not been carried out on the data set previously. In addition, though clustering is normally used to group objects or *cases* in a data set, this new analysis concentrated on the grouping of measured *variables* (i.e. capability and health variables) to examine the structure of the data set and to elicit the natural groupings of capabilities and disease conditions.

5.6.2.1 Clustering Procedure

Cluster analysis takes place in a series of stages. In the first stage, measures of proximity (referred to as similarity, dissimilarity or distance) between objects or variables are derived from the original data matrix **X**. This results in an *n x n* matrix of proximity measures for each pair of objects or variables, where *n* is the number of objects or variables being clustered. There are various measures of proximity that can be used depending on the type of data being considered (continuous, categorical or mixed). For example, dissimilarity measures for continuous data include the *Euclidean distance*, *City block distance*, *Minkowski distance* and *Pearson correlation* (Everitt et al., 2001).

Once the proximity measure has been selected by the researcher and the proximity matrix has been calculated, the second stage involves the selecting the clustering algorithm. There are different approaches to clustering including hierarchical clustering, non-hierarchical clustering, agglomerative clustering and divisive clustering. The most common approach is hierarchical clustering, where the objects or variables are classified in a series of partitions from a single large cluster containing all objects, to *n* clusters containing a single object. Agglomerative methods are the most widely used of the hierarchical clustering approaches, and the procedure consists of starting with *n* single member clusters and ending with a single cluster containing all *n* individuals (Everitt et al., 2001).

Agglomerative clustering algorithms work by fusing objects or variables that are most similar (determined by their proximity) into groups. There are different clustering algorithms for this process, as there are different ways of defining proximity between an object and a group containing several objects or between two groups of objects. Common clustering methods include *between groups linkage*, *within-groups linkage*, *nearest neighbour*, *furthest neighbour*, *centriod*, *median* and *Ward's method* (Everitt et al., 2001). As objects or variables are clustered into successively larger groups, these methods determine the way in which other objects or group of objects in the set combine to make higher level groups. Further mathematical details of each method are given in Everitt (2001).

After the researcher selects the clustering algorithm, the third stage involves looking at the output of the clustering in the form of a diagrammatic hierarchical tree structure known as a *dendrogram*. It is a mathematical and pictorial representation of the entire clustering procedure. Other output is provided including agglomeration schedules and icicle plots which give information on the clusters formed at each stage in the process. The final stage consists of interpreting the output and using the resulting clusters for the application at hand.

The dendrogram (see Figure 5-8) consists of objects or variables on the left side of the diagram and a line representing proximity or distance across the top. The diagram is read from left to right in the direction of increasing distance. The nodes of the dendrogram tree structure represent clusters, while the lengths of the stems represent the distances where clusters are joined. As the diagram is read left to right, the nodes and distances where clusters are formed become apparent, and the entire grouping structure can be seen at a glance.

The 7 capability and 15 health variables were clustered in SPSS data analysis software using all the agglomerative clustering algorithms available (between groups linkage, within-groups linkage, nearest neighbour, furthest neighbour, centriod, median and Ward's method). The output was then collated in the form of clustering dendrograms. Split sample validation was performed on the data set by selecting 50% of all cases randomly and re-running the clustering analyses. A similar structure was found for the split sample dendrograms compared to the full sample, providing evidence that the data set contains a consistent underlying structure.

5.6.2.2 Clustering disability scales

For the capability (or disability) scales, a Euclidean distance measure was selected for the proximity measure, and Ward's method was found to be the most effective clustering algorithm. Figure 5-8 shows the clustering output for the seven disability scales. The proximity measure used gives the 'distance' between every pair of the seven disability variables, and could be interpreted as the physical distance between two multidimensional points. Therefore, in looking at the proximity matrix, the smaller the distance between any two variables, the closer they are in terms of the distance analogy. Because the scales measure the *level* of disability in the population given by the score on the scale, the distances represent how similar the two variables are in terms of the level of disability across the population.

For example, the distance between seeing and communication has the lowest distance (156.375). This means that across the disabled population, the levels of seeing and communication disabilities are relatively close to each other. The communication and locomotion scales have the greatest distance between them (344.641) which means that the levels of communication and locomotion disability vary more widely across the population.

The dendrogram shows that seeing and communication first group together, followed by hearing. Intellectual functioning and reaching/stretching are then joined to this group at a further distance. Locomotion and dexterity form their own group at a larger distance. This structure of the data indicates that levels of seeing and communication are very close together across the disabled population. Hearing, intellectual functioning and reaching/stretching disability levels have somewhat larger variation in levels across the population. However, the largest variation occurs between levels of locomotion and dexterity indicated by the large distances at which these two variables group.

Proximity Matrix

Case	Matrix File Input						
	loc	reach	dex	see	hear	comm	int
loc	.000	299.436	291.830	336.941	335.584	344.641	328.317
reach	299.436	.000	256.517	198.553	216.006	181.648	219.201
dex	291.830	256.517	.000	306.398	309.495	305.543	303.310
see	336.941	198.553	306.398	.000	194.888	156.375	205.065
hear	335.584	216.006	309.495	194.888	.000	177.628	217.250
comm	344.641	181.648	305.543	156.375	177.628	.000	170.238
int	328.317	219.201	303.310	205.065	217.250	170.238	.000

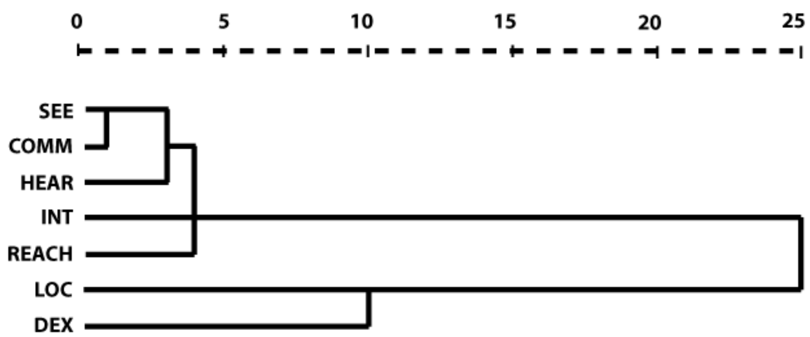


Figure 5-8 Proximity matrix and dendrogram showing clusters of capability loss (based on similar levels of capability loss in the population)

The implication if this is that across the disabled population, levels of seeing, communication, hearing, intellectual functioning and reaching disability tend to be similar. In other words, not only do these disabilities tend to co-occur, but also they tend to co-occur at a similar level or severity. Locomotion and dexterity disabilities tend to co-occur as well, however they manifest with greater differences in severity.

5.6.2.3 Clustering categories of disability

To further investigate the co-occurrence of disabilities in the population, the 7 capability (or disability) scale variables were converted to categorical disability variables i.e. the severity score value for each scale was replaced with a '1' if it contained a non zero value, or a zero otherwise. This procedure in essence created profiles for each case in the data set showing if a person had a disability or not across the 7 areas of disability. A Jaccard measure of similarity was used, and complete linkage (furthest neighbour) was used as the clustering algorithm.

Figure 5-9 shows groupings for the variables representing similarity in terms of co-occurrence. In this case, larger distance measures in the proximity matrix represent greater co-occurrence between two variables. It is evident from the dendrogram that that locomotion and dexterity disability are the most likely to co-occur. Reaching disability groups with these two variables to form a *motor capability group* at a greater distance. Hearing, intellectual functioning and seeing also group together forming a *sensory-cognitive group*. Communication is last to cluster with the previous two clusters due to lower levels of co-occurrence with the other groups.

Proximity Matrix

	Matrix File Input						
Case	cat_loc	cat_reach	cat_dex	cat_see	cat_hear	cat_comm	cat_int
cat_loc	1.000	.239	.385	.228	.290	.056	.258
cat_reach	.239	1.000	.362	.173	.166	.079	.194
cat_dex	.385	.362	1.000	.223	.236	.068	.233
cat_see	.228	.173	.223	1.000	.213	.082	.203
cat_hear	.290	.166	.236	.213	1.000	.077	.220
cat_comm	.056	.079	.068	.082	.077	1.000	.146
cat_int	.258	.194	.233	.203	.220	.146	1.000

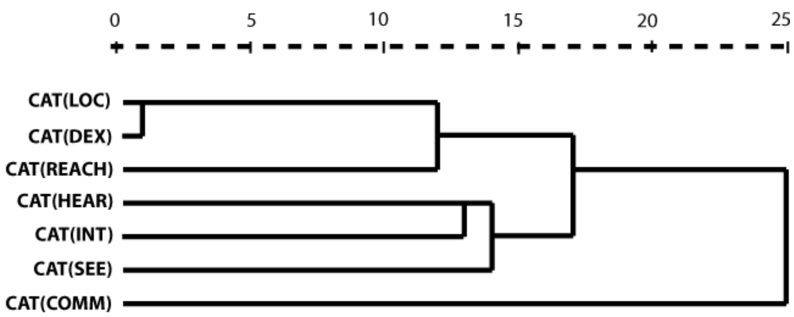


Figure 5-9 Proximity matrix and dendrogram showing clusters of capability loss (based on co-occurrence)

5.6.2.4 Clustering Health Conditions

To investigate co-occurrence of health conditions, the 15 disease category variables were clustered using a Jaccard measure of similarity and complete linkage (furthest neighbour) as the clustering algorithm (similar to the clustering of the categories of disability). Figure 5-10 shows groupings of medical conditions. Circulatory and musculo-skeletal conditions group at a very small distance, indicating that of all the medical conditions, these two tend to co-occur most frequently. Respiratory conditions tend to co-occur with circulatory and musculo-Skeletal diseases as well. Other co-occurring groups include eye and ear conditions, mental and nervous conditions, and neoplasms and genito-urinary conditions.

Proximity Matrix

	Matrix File Input														
Case	INFECT	NEOPLAS	ENDOC	BLOOD	MENTAL	NERV	EYE	EAR	CIRC	RESP	DIGEST	GENITO	SKIN	MUSCULO	OTHER
INFECT	1.000	.000	.005	.000	.007	.003	.004	.002	.003	.001	.003	.000	.006	.003	.010
NEOPLAS	.000	1.000	.022	.007	.010	.022	.022	.011	.029	.018	.028	.028	.006	.024	.000
ENDOC	.005	.022	1.000	.020	.035	.052	.065	.043	.137	.049	.048	.039	.031	.077	.006
BLOOD	.000	.007	.020	1.000	.008	.009	.021	.014	.017	.019	.010	.026	.010	.012	.006
MENTAL	.007	.010	.035	.008	1.000	.053	.014	.026	.036	.038	.032	.031	.018	.050	.015
NERV	.003	.022	.052	.009	.053	1.000	.051	.034	.077	.050	.048	.024	.012	.086	.006
EYE	.004	.022	.065	.021	.014	.051	1.000	.110	.070	.044	.035	.036	.014	.069	.013
EAR	.002	.011	.043	.014	.026	.034	.110	1.000	.062	.047	.036	.027	.018	.064	.010
CIRC	.003	.029	.137	.017	.036	.077	.070	.062	1.000	.102	.094	.046	.018	.202	.009
RESP	.001	.018	.049	.019	.038	.050	.044	.047	.102	1.000	.068	.030	.029	.109	.012
DIGEST	.003	.028	.048	.010	.032	.048	.035	.036	.094	.068	1.000	.051	.017	.102	.015
GENITO	.000	.028	.039	.026	.031	.024	.036	.027	.046	.030	.051	1.000	.008	.034	.003
SKIN	.006	.006	.031	.010	.018	.012	.014	.018	.018	.029	.017	.008	1.000	.021	.000
MUSCULO	.003	.024	.077	.012	.050	.086	.069	.064	.202	.109	.102	.034	.021	1.000	.009
OTHER	.010	.000	.006	.006	.015	.006	.013	.010	.009	.012	.015	.003	.000	.009	1.000

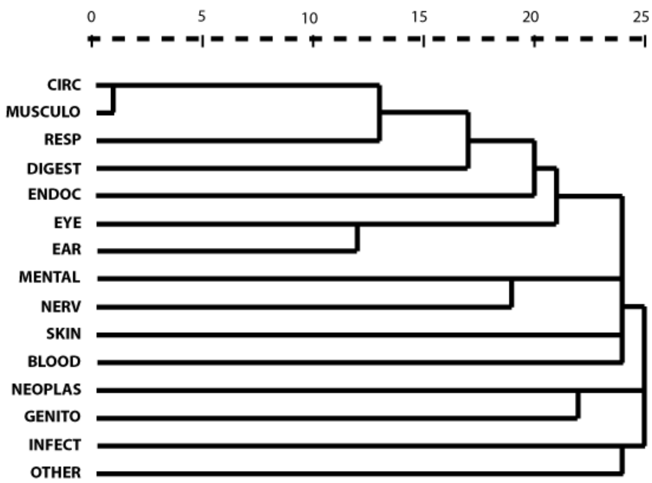


Figure 5-10 Proximity matrix and dendrogram showing clusters of health conditions (based on co-occurrence)

The groups generated by the clustering procedures were reliably structured but difficult to interpret, and it was decided to show the clustering dendrograms to two medical practitioners. A medical doctor was approached at the Newnham Walk Surgery and a Gerontologist was approached at Addenbrookes Hospital (both in Cambridge) for their interpretation. After explaining how the clustering procedure worked and allowing them some time to study the diagrams, two main issues emerged. Firstly, they experienced difficulty in interpreting the cluster groupings directly from a dendrogram due to unfamiliarity with the representation.

Secondly, they found it difficult to give underlying reasons for clustering from broad categories of health (as opposed to specific diseases/conditions). It appears that due to medical training, medical practitioners are skilled at making diagnoses based on specific symptoms, and relating these to various individual diseases. But at a higher categorical disease level, it was more difficult to make interpretations and find reasons that link the capability category to the disease category. Possibly any number of factors could link a family of diseases to various capability losses either directly or indirectly. They indicated that the diagrams, though interesting, highlighted the fact that associating capability groupings with medical conditions is difficult. However, they stated that if these relationships could be better understood, then such information could benefit not only designers, but also medical practitioners.

5.6.3 Discussion

Understanding the capabilities of the wider disabled population and how it is structured with age is an important first step to designing inclusively. Only after understanding the loss of capability that occurs with age and disabling conditions can designers begin to adjust products and environments that cater to such losses. The following insights are gained from the secondary data analysis of the DFS data.

Older and disabled people experience multiple capability loss caused by multiple co-occurring health conditions (Keates & Clarkson, 2003a). Inclusive Design requires the simultaneous consideration of motor, sensory and cognitive capability loss, as these losses tend to co-occur in older and disabled populations. In addition, a better understanding of combinations of sensory, cognitive and motor capability loss in relation to interface demands is needed for the successful design of more inclusive products. The patterns of co-occurrence differ with various age groups (Keates & Clarkson, 2003a), requiring designers to be aware of these patterns and how this impacts their product designs. This information could only be extracted from population representative capability data via multivariate analysis techniques (Carlsson et al., 2002). To date, such a design relevant data source does not yet exist (Johnson et al., 2010; Kondraske, 2006c).

Though it is difficult to associate levels of functional capacity with groups of medical conditions, information on the relationship between disease and capability could be useful in design. Understanding the lifestyle issues of people with various medical conditions and how these conditions could affect the ecology of product usage could lead to more Inclusive Designs. If broad groups of capability and medical conditions could be identified and related, this might result in a useful resource for both designers and medical practitioners. The ultimate goal would be a complete characterisation of users for design, with information on physical, sensory and cognitive capability issues being considered along with social and emotional considerations. More importantly, medical theory and general knowledge allows for linking and clarifying functional capability losses, particularly that identified in the clustering structure of the DFS data.

The high prevalence of musculo-skeletal conditions (such as all forms of arthritis that occur mostly with increasing age) and the loss of locomotion and dexterity capability suggest the design of products with reduced physical demand. The forces required to open packaging, operate various controls and physically manipulate products need to be reduced. The high prevalence of sensory ability loss in the older population requires that products provide stronger and clearer sensory signals (e.g. larger text, increased contrast, better lighting and adjustable volume levels). The prevalence of cognitive capability loss was shown to be greater in the younger age groups, possibly due to differential mortality in the older age groups (i.e. people with various conditions such as dementia do not survive, leaving the 81-100yr population with survivors). In any event, loss of fluid cognitive ability is known to decline with increasing age (Fisk et al., 2004), and product design must accommodate this.

The loss of ability to interact with everyday products generally tends to increase with increasing age. This is likely to be due to the effects of the ageing process and higher incidence of medical conditions that cause impairment leading to loss of ability. Designers are faced with meeting the challenge of designing inclusive products, services and environments that place minimum demands on people's sensory, cognitive and motor abilities.

The results of this exploratory cluster analysis study provide evidence for a coherent underlying structure in the DFS data. Cluster analysis also proved to be a useful tool for visualising the structure of capability and health in the DFS data, and similar multivariate methods might be useful in investigating similar data sets on health, disability and functional capacity. Given that there is structure and natural groupings of capability loss in older and disabled populations, it implies that product assessment methods should be designed to take this into account. Products should be evaluated with users that demonstrate this multiple capability loss, and even more importantly, the effects of multiple capability loss profiles on the ability to used products require further investigation. Though the Inclusive Design Cube

(IDC) model of capability provides a useable and simplified representation of capability loss in the population, it masks the true picture of disability being a non-linear, discontinuous space of dependent factors. Understanding this picture is critical to developing valid and robust product assessment methods.

5.7 Chapter Summary

In this chapter, the essential user capabilities for product interaction were described together with their changes with ageing and disability. The structure of disability prevalence in the UK population was also investigated via secondary data analysis of the 1996/97 Great Britain Disability Follow-up Survey (DFS) data. A complex picture of disability emerged where there are multiple capability losses across the population that need to be taken into account in product assessment. This chapter lays the groundwork for the following two chapters which explore the relationships between user capabilities and product demands in more detail.

Chapter 6 **Exploring Inclusive Interaction**

6.1 Chapter Overview

Based on the analysis of user capabilities presented in Chapter 5, inclusive interaction is explored in detail via empirical and analytical evaluations of a consumer product. Given the evidence for multiple interacting capability losses across the population, an observational study is first presented where two toaster designs were used by seven users with various multiple capability losses. The problems uncovered by users were recorded with an eye to further understand how capability loss impacts real-world interaction problems, and how the design of the product interface contributes to these problems. Secondly, a product demand analysis is presented for one of the toasters which demonstrates how interaction problems can be uncovered by the systematic consideration of product demands on user capabilities. The results of both methods are compared in terms of the type of problems found and how they can contribute to designing a more inclusive product.

6.2 Empirical Study with Two Toasters

The study described in this chapter was designed to fulfill various aims. Firstly, it was conducted to develop a better understanding of the types of interaction problems experienced by users with functional capability losses, and how these problems could arise from the mismatch between user capabilities and the demands of product features. Secondly, it was designed to compare an empirical method of product assessment (user observation) with an analytical method of product assessment. Thirdly, the study gave the author valuable practical experience and skills in working directly with older and disabled people as a precursor to further studies in the research process.

6.2.1 Empirical Study Design

For the empirical part of the study, two toasters were evaluated by seven participants. Toasters were chosen because they are fairly common consumer products with familiar interface features. It is also a product that supports a basic activity of daily living (eating/feeding) found in many kitchens in the home. Other kitchen products such as kettles have been investigated in depth (Cardoso, 2005; Keates & Clarkson, 2003a), and it was felt that toasters would make a suitable product category for further study. The two toaster designs used are shown in Figure 6-1. Toaster 1 had a relatively simpler design with three controls: a slider, a rotary heat control and a stop button. Toaster 2 had more features including a digital LCD display, memory functions and an assortment of buttons for setting the heat and performing other specialised actions.

A.Toaster 1 B. Toaster 2

Figure 6-1 Two toaster designs used in the empirical study

As the study did not strive for statistical generalisation, participants were sampled using a combination of convenience, purposive, snowball and heterogeneous sampling (Robson, 2002, p. 265). For an in depth qualitative understanding of issues, the literature suggests and provides guidance on working with a small groups of users (Cardoso, 2005; Cassim & Dong, 2003; Gheerawo & Lebbon, 2002; Kanis & Arisz, 2000; Porter, marshall, Sims, Gyi, & Case, 2003). In addition, study participants were selected based on the expectation that they would be able to use the product in question i.e. the product should not have been completely inaccessible to them. However, the product was selected to be challenging enough so that users would encounter difficulties with certain product features (Cardoso, 2005).

The main aim was to have participants with one or more impairments of vision, hearing, cognition and motor function that represent some of the profiles of multiple capability loss discovered in the data analysis of Chapter 5. The intent was also to have a few older users in the sample who might have minor memory loss that accompanies ageing. However, users with severe cognitive capability loss were not sampled for this study. Participants were recruited through personal contacts, the University of Cambridge Disability Resource Centre, the CAMSIGHT charity, and the Cambridge Engineering Design Centre database. All participants were volunteers and hence were not remunerated. The final seven participants were selected based on the preceding requirements and their availability. See Table 6-2 for further details on participant characteristics.

Given the four main interaction system components of the user, the product, the environmental context and the tasks to be performed (as explained in Chapter 1), it was decided to utilise the same 2 toaster products with the same tasks in different environments with different users. It was decided not to use one controlled environment for the product evaluations as the effect of the real-world context would be missed. It has been shown that user observations in real environments provide more accurate data on product usage (Leonard et al., 2006; Nielsen, 1993; Poulson et al., 1996; Robson, 2002; Rubin & Chisnell, 2008). In addition, since the study was qualitative in nature, strict experimental control was not required, and more was gained from observing the interactions in context (Leonard et al., 2006). The study was conducted at five different kitchen environments: (1) the Engineering Design Centre Loft kitchen (Participants 1 and 2); (2) the disability resource centre kitchen (Participant 3); (3) the Cambridge Faculty of English kitchen (Participant 4); (4) the Cambridge Faculty of English kitchen (Participant 5); and (5) the CAMSIGHT kitchen area (Participants 6 and 7). The study was conducted at a time convenient to the participants.

6.2.1.1 Empirical Study Procedure

The following details the study procedure followed:

Each participant signed a consent form at the beginning of the study. The consent form is given in *Appendix 1: Consent Form for Toaster Study*. The nature of the study and the reasons for conducting it were explained to each participant. It was also explained that the study was a test of the products and not a test of the participant. Each participant also filled out a brief questionnaire with general, health and product experience information. This questionnaire is given in *Appendix 2: Participant Questionnaire for Toasters Study*.

Each participant performed the task of making toast with Toaster 1 and Toaster 2 while being recorded with a Sony MiniDV video camcorder. The procedure used in Table 6-1 was followed. Participants were asked to voice their problems in a think aloud protocol as they performed the tasks (Boren & Ramey, 2000). While participants were speaking, there was minimal input from the researcher. The researcher acknowledged by responding with tokens such as ‘mm hmm’ or ‘uh huh’ to encourage the participant to continue talking without introducing any bias (Boren & Ramey, 2000). If necessary, non-leading probing questions were sometimes used, for example, “how easy or difficulty is it to perform the task?”

At the end of the session, participants were debriefed and thanked for their participation.

Table 6-1 Study Procedure Used By Researcher

Toaster 1 Set up toaster plugged in, but with power off. Set up bread bag and bagel bag closed (so participant has to open it). Ask the participant to make two slices of toast with the toaster on setting 5. While toasting, ask participant to stop the toaster and pop toast out. **Toaster 2** Plug in toaster 2 and set it up for participant. Ask participant to make 2 slices of toast at setting 5. While it is toasting, ask them to stop it. Ask participant to make a bagel on setting 5. They have to cut the bagel. Stop the toaster while the bagel is toasting. ASK: How do you think the memory setting might work? Try to set the memory to whatever they like. Tell participant to let the bagel toast till the end. Ask participant for any other thoughts on the toaster (how easy it is to use, what they found particularly difficult etc.) ASK: Between these two toasters, which toaster do you like better? Which one would you buy?

6.2.2 Data Analysis

The interaction videos from each participant were transferred from miniDV tape to a digital video file format using Adobe Premiere Pro. Each video was reviewed first in its entirety and then reviewed again while recording problems encountered by participants. Following this, multiple passes were made through each video to ensure that all problems were recorded. Data from the questionnaire was logged in an excel spreadsheet. The results of the empirical observation are presented in the following section.

6.2.3 Results

The profiles of study participants are shown in Table 6-2. The age range of participants spanned 24 years to 64 years, with the mean age being 43years (SD=15.72). The sample consisted of four females and three males. The educational background of participants varied, with one participant having GCSE Ordinary Levels, three participants with a degree, and three participants with a postgraduate qualification. Participants reported various medical conditions which are given in Table 6-2.

In terms of problems with everyday activities, 4 participants reported problems with their vision and 2 participants reported problems with their hearing. Only one participant reported problems with memory. In terms of physical activity, 3 participants reported problems with lifting objects, 2 participants reported problems with walking, 2 participants reported problems with twisting and turning dials and knobs, and 2 participants reported problems with grasping, pulling or pushing. Participants also reported using various aids such as joint supports (splints), magnifiers and monoculars, a computer with larger print, canes, hearing aids, and a microlink system that reduces the background noise. There was one powered wheelchair user in the sample (participant 3). In the process of conducting the study, participants were allowed to use whatever aids they would normally use in the course of daily activities.

All participants reported that they currently owned a toaster, and 6 of the 7 participants had more than 5 years experience using toasters in general. One participant reported less than one year using a toaster (participant 5).

Table 6-2 User information with pictures of toaster interaction

User Information	Toaster 1	Toaster 2
Participant 1 Age 24, Female Conditions: Rheumatoid and Psoriatic Arthritis		
Participant 2 Age 30, Male Conditions: None		
Participant 3 Age 50, Female Conditions: Tetraplegic following spinal cord injury, Fulltime wheelchair user, paralysed from neck down with limited hand and arm function		
Participant 4 Age 64, Female Conditions: Osteoporosis, requires aids to read fine print		
Participant 5 Age 27, Male Conditions: Cortical visual impairment, nystagmus (inability to fix gaze), tunnel vision		
Participant 6 Age 51, Male Conditions: Stickler Syndrome, glaucoma, cataracts, degenerative retinas, Acoustic neuroma, deaf in one ear		
Participant 7 Age 55, Female Conditions: Deteriorating hearing and sight conditions		

6.2.4 User Problems: Toasting Bread with Toaster 1 and Toaster 2

The main problems encountered by participants when making toast with toaster 1 are given in Table 6-3 together with frequencies of occurrence among the sample. The top part of the table shows the number of participants reporting difficulties with various actions in their everyday life. 5 of the 7 participants found difficulty in reading the numbers and arrow on the heat control (problem 1). This was due to the placement of the control on the right hand side of the toaster, and the small size of the text and arrow symbol. Problem 2 of the second highest frequency (3 of 7 participants) was that of participants pressing the slider to make toast without making sure it was first on at the mains switch. When the slider was pressed, it did not stay down indicating that the toaster was unpowered. Only after realizing this did participants turn the toaster on at the mains. This indicated that the toaster did not sufficiently provide feedback on whether it was in a powered state or not.

Table 6-3 Problems encountered by participants using toaster 1

	Participants							
Difficulties with	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Freq**
Vision								4
Hearing								2
Remembering what you were doing (your memory)								1
Walking								2
Lifting objects								3
Twisting and turning dials and knobs								2
Grasping, pulling or pushing								2
Problems Encountered	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Freq**
1. Arrow mark and numbers on heat control were not adequately visible								5
2. Slider was pressed without first turning on power at the mains								3
3. Difficulty encountered in seeing plug and switch at the mains								2
4. Difficulty encountered in opening bread packaging								1
5. Difficulty encountered in reaching the plug switch at the mains								1
6. Toaster slid on counter surface while operating heat control								1
7. Light reflection caused difficulties in reading the text on the toaster								1
8. Lack of tactile support on the control caused difficulties in operation								1
9. Difficulty encountered in seeing stop button								1
10. Difficulty encountered in seeing toaster chassis against background environment								1
11. Difficulty in reading the ‘stop’ button text								1
12. Difficulty in seeing the toaster slider								1

Problem 3 of the third highest frequency (2 of 7 participants) was that of participants not being able to see the plug and switch at the mains due to low contrast (white on white). This caused problems for the participants with visual capability loss, highlighting the effect of environmental conditions on product usage. Other problems such as problems 8, 9, 10, 11 and 12 were caused by lack of contrast, adequate size or environmental conditions making visual tasks difficult. Participant 3 with relatively low hand function had problems with the rotary control, though other users experienced no difficulty with manipulating the control. Participant 3 also had problems with reaching the plug and switch at the mains given that she was a wheelchair user (problems 4 and 5). Problems 6 and 7 were caused by environmental conditions such as the smoothness of the kitchen counter surface and the ambient lighting conditions.

In general, participants encountered problems based on the type of capability loss experienced. For example participants 5, 6 and 7 with visual capability loss encountered problems with visual tasks such as reading, and environmental problems such as glare. Interestingly, these three participants checked to see whether the toaster was on at the mains before trying to operate the toaster and thus avoiding problem 2. This could be attributed to a learned behavior of people with visual capability loss, as they might not be able to easily detect some product state indicators due to poor vision. To compensate for this, they would make sure to check the mains power for a device before starting a task with a product.

In the case of participant 5 with visual capability loss, a problem with the tactile feel of the heating control was encountered. This participant was unable to see the heating control clearly and therefore relied on tactile feedback to operate it. Because the control was smooth, he experienced difficulty in turning it to the desired setting. The control was also a continuous control, and there were not discrete stops or ‘clicks’ that could assist in setting it at the right value. This serves as an example of one sensory modality being used to compensate for a lack of another, representing one form of a coping strategy.

In other cases, for example with participant 1, though motor capability difficulties in daily life were reported, there were no motor difficulties encountered in the product task. Therefore it can be concluded that the participant possessed sufficient motor capability to perform all motor tasks required of the product. Without knowing the specific level of capability loss of the user and the specific capability demand of the product, it is not possible to determine if a user can or cannot perform a task. This implies that knowing whether people experience capability difficulties in daily life is not enough to make judgments about the difficulty experienced in product usage.

Table 6-4 lists the main problems encountered by participants when performing the task of making toast with toaster 2. Problem 1 resulted in confusion over the function of the small sliders on toaster (used for ejecting the toast out of the slots). Problems 2, 3, 6 and 7 involved difficulties in seeing and reading interface features. Problems 4, 5 and 7 resulted from misunderstandings about the mapping of interface features to each other, and also in understanding what certain features and symbols meant.

Table 6-4 Problems encountered by participants using toaster 2

	Participants							
Difficulties with	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Freq**
Vision				■	■	■	■	4
Hearing						■	■	2
Remembering what you were doing (your memory)				■				1
Walking	■		■					2
Lifting objects	■		■	■				3
Twisting and turning dials and knobs	■		■					2
Grasping, pulling or pushing	■		■					2
Problems Encountered	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Freq**
1. Difficulty encountered in understanding the function of the small levers (used to eject toast)	■	■		■				3
2. Difficulty encountered in reading the LCD display					■	■		2
3. Difficulty encountered in reading text on buttons					■	■		2
4. Erroneously placed toast in two middle slots rather than using a pair of slots on the ends				■				1
5. Found symbols on buttons confusing				■				1
6. Difficulty encountered in reading the label on the slider					■			1
7. Unsure what buttons do					■			1
8. Light reflection caused glare problems							■	1

The usage of toaster 2 resulted in a smaller number of problems overall than toaster 1. This was due to certain visual problems being avoided with toaster 2. For example, since the plug and cord were black on toaster 2, there was high contrast against the white mains fixture. The toaster was also larger and heavier, so it did not slide on the kitchen counter surface. All interface features were to the front of the toaster for easy access, and it indicated the 'ON' state readily as the LCD display came on once there was power available. However, due to the increased number of interface features on toaster 2 such as buttons with icons and a screen, more cognitive problems were introduced compared to toaster 1.

As with the case of toaster 1, participants encountered problems based on the type of capability loss experienced. For example participants 5, 6 and 7 with visual capability loss all experienced different types of visual problems. Participants 1, 3 and 4 with motor capability problems in daily life did not experience any significant motor problems with toaster 2, implying that their motor capability was above that required to use the toaster. Participant 4 who reported problems with memory in daily life encountered the most problems with understanding the features and functions of the toaster.

6.2.5 User Problems: Toasting a Bagel with Toaster 2

Participants were asked to toast a bagel with toaster 2 in order to observe if they would use the special ‘bagel’ function button that was provided on the interface. Table 6-5 lists the problems encountered. Two participants did not use the special ‘bagel’ button provided in toasting the bagel. Two participants also encountered problems with the lifting lever that lifts the bagel up out of the slots. The lever requires some additional force to be depressed which was found difficult and unresponsive by participants 2 and 7. Other problems existed in terms of the symmetry of the interface and mapping the bagel button to the pairs of slots on either side of the toaster. Users expected the bagel and crumpet button to be replicated on the right side of the toaster as was done for some other controls e.g. the cancel/stop button (see Figure 6-2). One participant did not attempt to use the lifting lever at all, and another complained of the lack of tactile feedback on the shiny top surface of the toaster. The shiny surface caused glare problems in the kitchen environment for one user with visual capability loss.

Table 6-5 Problems encountered making a bagel with toaster 2

	Participants							
Difficulties with	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Freq**
Vision				■	■	■	■	4
Hearing						■	■	2
Remembering what you were doing (your memory)				■				1
Walking	■		■					2
Lifting objects	■		■	■				3
Twisting and turning dials and knobs	■		■					2
Grasping, pulling or pushing	■		■					2
Problems Encountered	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Frequency**
Did not use the bagel button to toast a bagel	■						■	2
Difficulty encountered in using the lifting lever		■					■	2
Difficulty mapping bagel buttons to either side of toaster					■			1
Removed bagel without using the lifting lever					■			1
No tactile feedback near the top of toaster						■		1
Shiny steel surface causes glare							■	1

6.2.6 User Problems: Setting the Memory on Toaster 2

Participants were asked to figure out how the memory function works on toaster 2. Figure 6-2 shows the control layout of toaster 2, with the 'Memo' button on the left. The toaster provides four different memory programs that cycle through when the 'Memo' button is pressed i.e. a number from 1 to 4 is displayed on the screen. To actually set the memory, the 'Memo' button is held down for a few seconds and the program number will flash on the screen. The user can then select various functions such as toasting level, bagel, crumpet etc. and then press the 'Memo' button once more to set the memory. The program number will stop flashing when this is done.

Figure 6-2 The interface of toaster 2 consisting of buttons, sliders and an LCD screen

No participant in the study was able to successfully set the program memory on the toaster. Participants tried to understand how it worked by using different strategies, but all ended up utilising trial and error behaviour where they resorted to random button presses when their strategies did not work. The key step was figuring out that the 'Memo' button had to be depressed for a few seconds before any memory program could be set (as indicated by the flashing program number). However, participants could not go beyond cycling through the memory program numbers. All participants questioned the use of such a memory function on a toaster after finding difficulty with the task.

6.2.7 User Preference between Toaster 1 and Toaster 2

Participants were asked to indicate their preference between toaster 1 and toaster 2. Five of the seven participants indicated that they preferred toaster 1, while the remaining two participants indicated that they did not find either of the two toasters appealing. The advantages of toaster 1 included it being more direct in its controls, simple, compact, lightweight (easy to lift and put away) and having no electronic displays. In terms of disadvantages, one participant commented that she would “not buy a toaster like toaster 1 because it has no contrast (as it is white) and you have to think of the kitchen décor”.

Three participants commented that toaster 2 ‘looked nicer’ than toaster 1. In addition, toaster 2 had various advantages including big slots to fit crumpets and bagels, a nice shape and feel, nice buttons, crumpet and the bagel features, and all controls are located on the front of the toaster. The disadvantages of toaster 2 that were mentioned included no real need for a memo function, confusing and complicated interface, large size, and poor display visibility due to ambient lighting.

6.2.8 Issues Observed

Taking into account the observed problems in using the two toasters, the following general interaction issues were extracted:

1. Adaptability and coping strategies: There was a difference between assumed ‘standard’ interaction processes and the reality of real interaction when people with disabilities used the toasters. Given sufficient motivation and people’s capability to adapt, users tried to accomplish a task using unexpected or non-standard actions (Kanis, 1993). For example, users used different parts of the hand or parts of the body for force exertion when they lacked the capability to perform the assumed standard manipulation (Kanis, 1993). Participant 3 was able to rotate the heating control on toaster 1 even though she could not form the pinching grasp required for rotation. The rotational force required by the design was very low, and the control afforded the use of one finger to gradually rotate the control in place.

Another example involved participants with visual capability loss moving in much closer to the product and adopting kneeling or stooped positions in order to see interface features. Participant 5 in particular mentioned the ability to make modifications to the product by adding Braille or tactile feedback where he wanted. Though the goal of Inclusive Design is to make the product as inclusive as possible at the outset, there may be value in designing products to be ‘hackable’ or customisable at the user’s discretion. This would allow for further personal modifications to the product that may be difficult to anticipate in the design process when catering for people using heterogeneous coping strategies.

In a recent study by Yoxall et al. into to openability of packaging, it was found that elderly people resort to both physical and social coping strategies (Yoxall et al., 2010c). Physical strategies can involve the use of modified actions as found in this study, or the use of improvised tools e.g. knives and cloths used to open a jar. Social coping strategies involve asking family members, friends or even acquaintances for assistance in performing the task. Women also utilise a larger number of coping strategies at an earlier age compared to men (Yoxall et al., 2010c). Importantly then, coping strategies extend beyond the narrow scope of direct product interaction into the larger physical, social and emotional usage environment. This highlights the importance of both the physical and social *context of use*. This is an area ripe for further research and development.

2. Sensory-cognitive interaction: It was found that expectations directed users with sensory capability loss in finding product features. Even though some users could not clearly see the slider on the side of toaster 1 due to low contrast, they were still able to find it because it was expected to be located on the side of the toaster. In this regard, experience/knowledge stored in long term memory drove sensory perception. Utilising these accepted interface norms could make locating interface features easier for users with sensory capability loss. As sensory input is the gateway to further cognitive processing, once the sensory input is degraded, effective cognitive processing becomes difficult. Thus innovating with new interface features (controls and displays) poses a challenge for such users who might rely on accepted interface norms for successful interaction.

Sensory-cognitive interaction theories essentially postulate that sensory functioning can directly and interactively affect cognitive functioning as in the case of older drivers (Baldwin, 2002). The cognitive workload increases when there is a degradation of sensory information. Given that sensory and cognitive functions are so closely linked and dependent upon each other, this becomes a critical area for further research and development. The evidence of the disability structure in Chapter 5 showed that sensory and cognitive capability loss tend to co-occur at roughly similar levels of severity in the population. Therefore understanding sensory-cognitive interaction becomes a critical area for consideration in terms of product assessment for Inclusive Design.

3. Learning and trial and error behaviour: Participants tended to resort to trial and error behaviour when they encountered an unfamiliar situation (i.e. they did not know how to use a particular feature or they became 'stuck' at a point in the interaction and could not figure out how to proceed). This form of behaviour could be attributed to reinforcement learning where actions that appear to take the user closer to a goal are learnt (Langdon, Lewis, & Clarkson, 2007). The setting of the memory on toaster 2 demonstrated this effect. The interface was not designed with any obvious cues of how to set the memory, nor any support for easily learning how to perform this task. Though the manual could have been consulted, the fact remains that the design of toaster 2 left much to be desired in terms of supporting the learnability of product features.

4. Aesthetics versus accessibility: Though the majority of the participants preferred toaster 1 over toaster 2 due to its small size and simplicity, three participants still commented on the appealing look (aesthetics) of toaster 2 and its fit into a kitchen environment. This trade-off between aesthetics and accessibility is an important consideration for designers, as it is a fine balance between designing for visual appeal as well as maintaining accessibility for various user groups. For example, the material selection of stainless steel that gives toaster 2 its appealing aesthetic caused problems with glare under certain lighting conditions. The challenge remains for designers to utilise appropriate material selection and form factors to achieve the desired aesthetic appeal, but in a way that maximises product accessibility. It could be interpreted as an avenue for innovation in product design and material selection.

6.2.9 Summary

Given these results and considerations, an analytical assessment of the product demands of toaster 1 is presented in the following section by examining the sensory, cognitive and motor demands of the toaster.

6.3 Product demand analysis

In this section, an analysis of product demands on user capabilities is presented for the simple toaster on the left of Figure 6-1. The toaster features two slots for bread, a rotary heating time control, a slider to activate the toaster and a stop button to eject the bread while toasting.

6.3.1 Sensory Demands

Figure 6-3 shows the simulated appearance of the toaster when viewed with loss of visual acuity, contrast sensitivity and peripheral field (created in Adobe Photoshop graphics software). In (b), the blurriness caused by lack of adequate acuity under operating conditions causes problems in reading the label on the stop button and the labels on the heating control. Loss of contrast sensitivity in (c) can make it difficult to distinguish the heat control or the slider from the toaster chassis. Finally in (d), loss of peripheral visual field can cause problems when trying to relate controls to product functions and to each other.

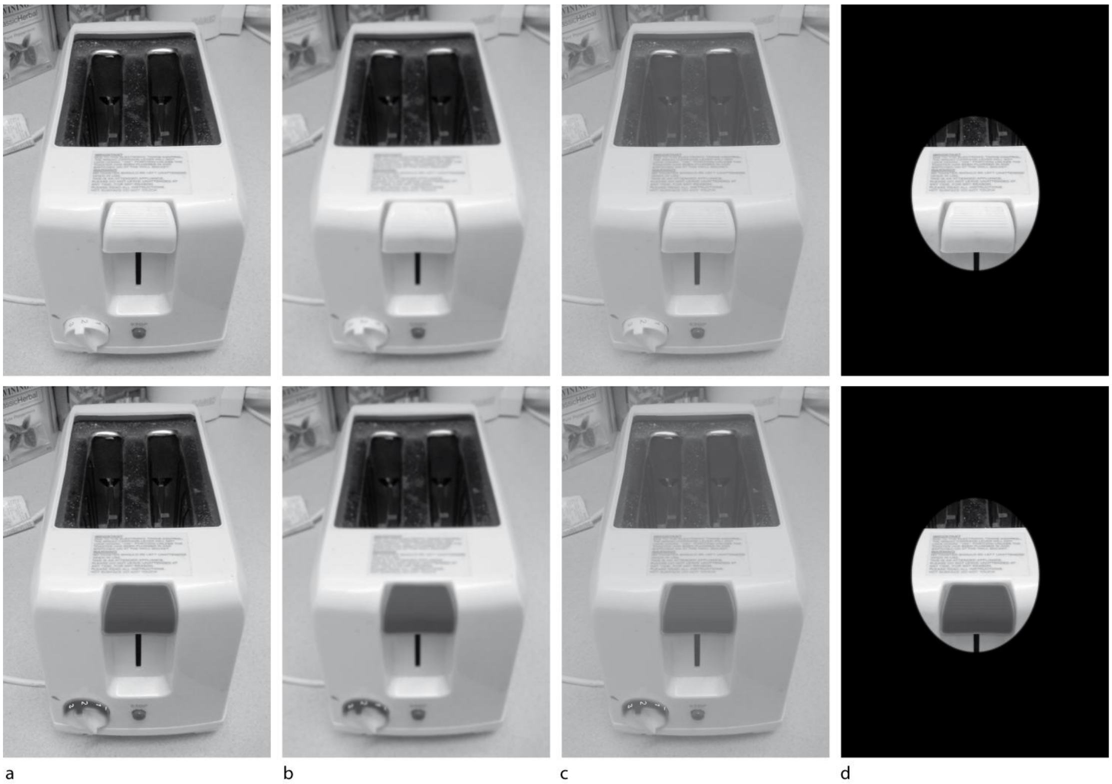


Figure 6-3 The top row (a) shows a simple toaster with (b) simulated loss of acuity, (c) simulated loss of contrast sensitivity and (d) simulated loss of peripheral visual field. The second row shows the same simulations with improved contrast on the heating control and slider.

Despite the simplicity of this particular toaster, increasing the contrast of the heating control and the slider against the toaster chassis simulated in the second row of Figure 6-3 can make them much easier to detect for people with reduced acuity, reduced contrast sensitivity and reduced visual field. Though the toaster does not have any explicit hearing demands, design improvements could be made by possibly including an audio alert when the toast is finished cooking to alert the user that toast is done. These changes suggest direct improvements to the toaster that would address problems 1, 3, 9, 10, 11 and 12 found in the user study that all had to do with small sizes and poor contrast.

6.3.2 Cognitive Demands

Figure 6-4 shows a representation for the reactive behaviour of the toaster using a visual formalism known as a state chart (Thimbleby, 2007). State charts are part of the Unified Modelling Language and thus provide a standardised way of representing the response of the toaster to various user actions. State charts are also commonly used to specify the design of reactive systems and are familiar to designers of embedded systems (Thimbleby, 2007). The demanded user action sequence for making toast is also shown in the figure consisting of six steps.

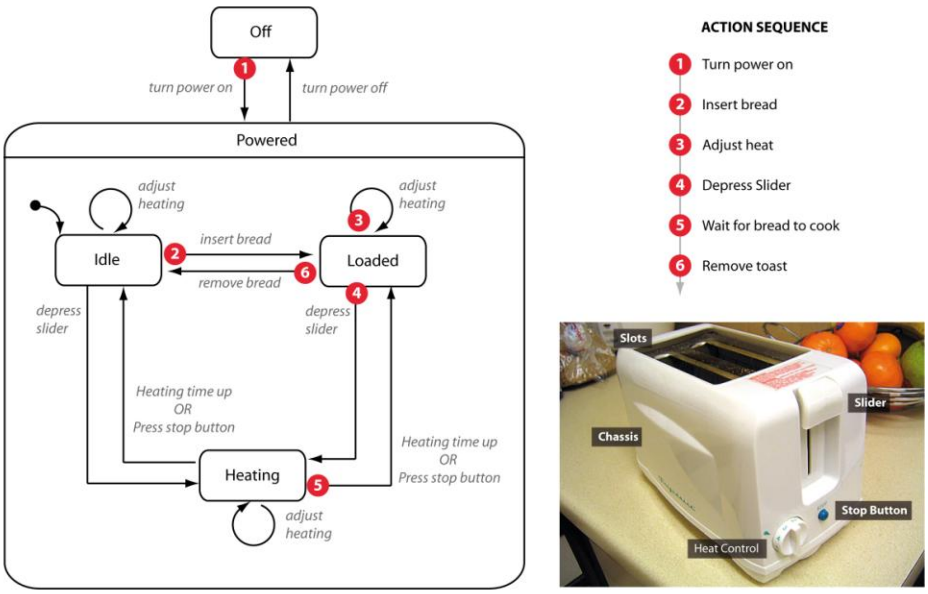


Figure 6-4 Diagram showing a state chart of toaster reactive behaviour (left) with a demanded action sequence overlaid, the demanded action sequence (top right), and a picture of the toaster with relevant interface features labelled (bottom right)

The representation can be used to derive the action sequences that form users' mental models by analysing variations to the demanded action sequence. For example, if step 1 of powering the toaster is omitted, an error of omission occurs and the toaster slider will not activate. In addition, the state based representation of the use process allows for the evaluation of adequate feedback on each state of the device. By examining the toaster, it is evident that inadequate feedback is provided in the powered state and in the heating state. This was the second most frequent problem (problem 2) found in the user study. An improvement to the design is immediately evident by including a visual signal such as an indicator light to indicate that the toaster is powered. An auditory feedback on reaching the bread cooked state would be another improvement to indicate to the user that the toast is finished, which can also benefit visually impaired users.

Different paths through the state-space of the toaster could be represented as shown in Figure 6-5. A state can be defined in terms of five variables: (1) the bread being cooked or uncooked, (2) the heat control being set or unset, (3) the slots being filled or empty (4) the power being on or off and (5) the slider being activated or inactivated. The values of the initial state variables are shown in the figure and the goal state is reached when a user has made toast. In general a user possibly has to keep track of three things at once: the state of the bread, the control setting and whether the toaster is cooking or not. If the assumption is made that most people leave their toaster on a particular setting most of the time, the number of items that needs to be tracked drops down to two. Therefore the toaster does not demand a high working memory capacity.

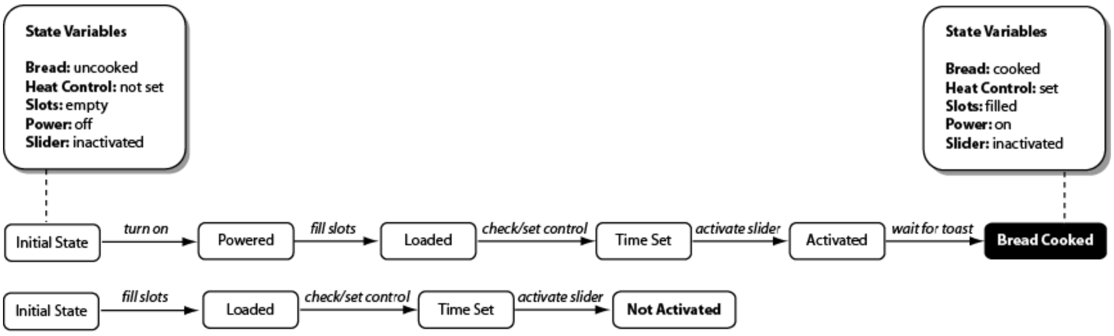


Figure 6-5 Two paths through the state space of the toaster with only one path leading to the goal

The state-action representation allows for the evaluation of feedback on each state of the device. This supports the user forming an adequate mental model of device operation where a good understanding of cause and effect actions can be learnt. Figure 6-5 also shows a second path through the state space of the toaster that results in no toast being made. In this path, the user forgets to plug in the toaster, and due to built in constraints, the toast does not activate. Again, a signal to indicate that the power is on would be an improvement to the current toaster design to remedy this.

If the assumption is made that people commonly leave their toaster on a particular heat setting most of the time and the power is on; the user has to plan for four steps: (1) put bread in, (2) depress slider, (3) wait for bread to cook and (4) remove toast. Thus the user is only required to keep track of the state of the bread. The attention demands are therefore relatively low assuming that there are no other distractions in the cooking environment that would be outside the control of the toaster designer. However, apart from the toast popping up, the toaster does not signal that the bread has been toasted. As previously mentioned, an auditory alert could be a design improvement, especially if users may attend to other tasks while bread is toasting.

The number of steps in the action sequence can be used as an indicator of the level of demand on planning capabilities of the user as he/she plans through the sequence of actions that will be performed on the toaster. There are also no time demands for task actions that could exceed the working memory time limit on storage. For any given design, the aim should be to reduce the number of actions that users have to perform in order to reach their goals thus reducing the cognitive demands of planning and working memory.

Toasters are relatively common consumer products, and the design of toaster 1 follows the traditional toaster form factor with slots at the top and a slider at the side. The interface features of slots, slider, rotary control and button are standard features that are also straightforward and prevalent on other toaster designs. Thus the demands of the toaster on the knowledge and long term memory of users is assumed to be relatively low. However, the toaster introduces a prospective memory demand where the user is required to remember to remove the toast once it is finished cooking. Again, the simple toaster does not signal that the bread has been toasted, or is being toasted, apart from the toast popping up.

Communication and language demands are relatively low for the toaster because no graphical symbols are used. The user is only required to read and understand the 'STOP' button label and the red text of a safety sticker on the toaster. Reading the safety sticker is not essential to using the toaster, so within the defined task bounds the toaster does not demand a high degree of written comprehension capability. Visuo-spatial communication demands are also relatively low for the toaster as no graphical symbols are used.

Feedback plays an important role in the learnability of a product. As previously mentioned, the representation of mental models allows for the evaluation of adequate feedback on each state of the device. This supports the user forming an adequate mental model of device operation where cause and effect associations can be learnt. When faced with novel products or situations with few generic features, users typically resort to trial and error exploration of the interface. By designing the product to support this type of exploration through salient feedback of its current state, users will be supported in their attempts to learn how the product works.

6.3.3 Motor Demands

The toaster can be analysed for its motor demands in a systematic way by examining the details of the task and extracting the necessary physical actions for interaction. Information about the type and number of contacting body parts (number of hands, finger, and palm), the action (e.g. lifting, pressing, grasping, turning) and directionality (linear in three principal dimensions, rotational) can be recorded in a tabular format as shown in Table 6-6. Figure 6-6 shows some of these dexterity demands. Most of the motor actions required to operate the toaster controls can be performed with one handed, non-compound actions. This results in a design that is accessible to one handed users. In order to evaluate these motor actions for exclusion, data is required on the maximum weight that can be lifted with each hand using a pinch grasp (a), the maximum torque that can be applied with each hand when turning clockwise using a two finger pinch grasp (b), the maximum downward linear force that can be exerted with fingers of each hand (c), and the maximum index finger push strength for each hand (d). Each control feature can be evaluated for exclusion based on the comparison of such data to the particular demand of the control.

Table 6-6 Physical demands of toaster

Main Tasks	Description
1. Setting heat control	1.1 Reach toast control, 1-handed action 1.2 Grasp control, pincer grasp 1.3 Rotate control to required position, rotational torque
2. Pressing slider	2.1 Reach to slider, 1-handed action 2.2 Depress slider till it clicks, 1-handed, linear downward force with fingers
3. Pressing stop button	3.2 Linear finger force
4. Removing toast from toaster	4.1 Reach toast, 1-handed action 4.2 Grasp toast, pincer grasp 4.3 Remove toast from toaster, vertical lifting force

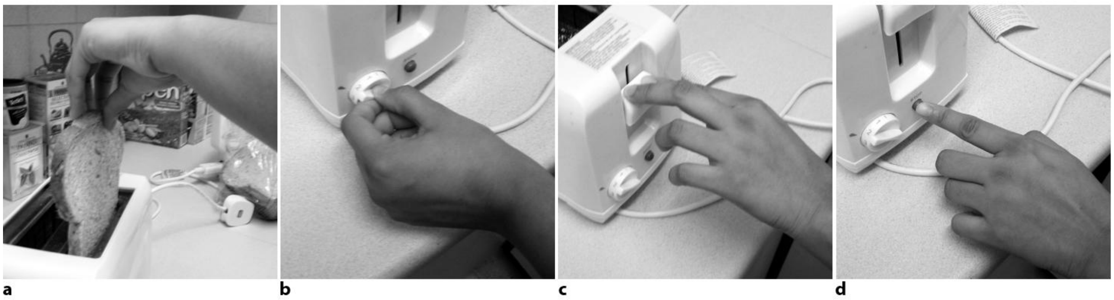


Figure 6-6 Dexterity demands of toaster controls (a) grasping bread (b) heat control (c) depressing slider (d) pressing stop button

Features can also be ranked in terms of importance to the task at hand. For toaster 1, almost all the control features described are essential to making toast. Other toasters come with features such as re-heat buttons, defrost buttons and memory functions (for example toaster 2 in the empirical study). These other features are not essential to the main task, and as such may not result in a large increase in motor exclusion if they are present. However, they may induce extra cognitive demands on the user.

As was discovered in the user study, users with capability limitations tend to employ various coping strategies when interacting with products that exceed their abilities (Kanis, 1993; Yoxall et al., 2006). For example they may use two hands instead of one or use body parts to exert forces on the product. However, because coping strategies can be quite varied, it is necessary to adopt a practical approach in assuming that users will want to interact in a normative way, and in analytical evaluation, these assumptions must be made. To supplement this, a series of coping strategies could be developed that would supplement these assumptions for different user groups. This remains an area for further research.

Problems with a lack of tactile support on controls (as was problem 8 in the user study) could be addressed as a standard check for each control or set of controls on the product. Other motor problems such as opening the bread packaging (problem 4), reaching the plug switch (problem 5) and the toaster sliding on the counter surface (problem 6) all depend on the environment and context of use. This could be addressed by developing a list of packaging types, expected distance ranges from product to plug and common surface types likely to be encountered. These will all have to be considered in the product design.

6.4 Discussion

This study consisted of both an empirical evaluation component and also an analytical evaluation component. The empirical study brought to light the richness of user interaction with consumer products and areas for design improvement. With a small sample size of seven users, some of the main problems with both toaster designs were highlighted. The empirical evaluation also provides for empathy on the part of the designer/researcher and it can help identify inspiration for innovations in product design. However, in such a small sample, it is unlikely that all the main user groups covering the breadth of sensory, cognitive and motor capability loss could be represented.

The empirical component of the study demonstrated that users generally experienced problems in the areas of their capability loss, but without more information on the specific level and severity of the losses, predictions cannot be made about whether a user will encounter a specific problem or not. In addition, the study highlighted the complexity of real world interaction and the effects of multiple capability loss. These important issues included the use of coping strategies, sensory-cognitive interaction, learning and trial and error behaviour for difficult tasks, and the aesthetics versus accessibility trade-off. The importance of understanding the context of use was also highlighted.

Conducting the empirical study gave the author valuable experience in working with older and disabled people as a precursor to the larger study presented in Chapter 7. It was found that older and disabled users require more time, patience and encouragement when performing usability evaluation studies. For users with sensory capability loss, it is important to provide study materials with large print and to speak clearly and slowly so that they could process the information being given.

The analytical evaluation performed by the author for toaster 1 demonstrated that the use of a systematic analysis coupled with tools such as statecharts can also flag important areas for design improvement. This included raising issues such as the contrast of the slider, the lack of a powered status indicator and the lack of an alert to indicate that toast is completed. However, the analytical method cannot easily address issues with coping strategies and user preferences. The analytical evaluation method would be further enhanced if it could provide quantitative estimates of how many people would be excluded by the design, and how many people would find it difficult to use in a given population. This would serve not only design functions, but also marketing functions as well. In order to provide such quantitative estimates, a user capability database would have to be available providing capability data for every sensory, cognitive and motor product demand. In order to make valid estimates, the relationship between pre-measured capability values and the product demands would need to be carefully modelled and understood. This issue would be explored with the study presented in the next chapter.

Both types of product evaluation methods lead to results that would be useful for product design, and both methods could be used in the Inclusive Design process. However, predictive analytical methods that rely on user capability data are less developed than empirical methods in the field of Inclusive Design. Thus, there is scope for further work in improving analytical evaluations with supporting user capability data.

6.5 Chapter Summary

After demonstrating the co-occurrence of multiple capability loss at the population level in Chapter 5, the aim of the preliminary study conducted in this chapter was to examine how capability loss impacts real-world interaction problems via an empirical study with toasters. The complexity of interacting capability loss became apparent with the observation of coping strategies, sensory-cognitive interaction, learning and trial and error behaviour, the trade-off between aesthetics and accessibility and the context of use.

An analytical evaluation of one toaster was carried out and the results showed that an analytical evaluation process could indeed find design problems that are validated when real users use the product. The analytic assessment can be used to improve products and therefore has its place among the toolkit of the designer. Further, a case was made for coupling the analytical evaluation process with a user capability database to make quantitative estimates of design exclusion. This would enable the designer to make design changes and improvements based on real data. The next chapter investigates the issue of describing and modelling the relationship between pre-measured capabilities and product demands in actual use situations as the next logical step.

Chapter 7 **Experiment: Four Consumer Products**

7.1 Chapter Overview

This chapter describes an experimental study designed to investigate the third research question: What relationships exist between measures of human functional capability and measures of task performance in the context of a user capability-product demand model of interaction? An empirical study was conducted using four consumer products used in activities of daily living: a clock-radio, a mobile phone, a blender and a vacuum cleaner. The study was designed to explore the interaction between users with multiple capability losses and multiple products drawn from different categories. Nineteen older and disabled users were recruited, and their sensory, cognitive and motor capabilities were measured using various capability tests. Participants were assigned to perform one task with each of the four products while being videotaped. After performing each task, difficulty ratings were collected for the main actions performed. Task times and errors were extracted from the resulting video data. The results were plotted and analysed to determine the relationships between user capabilities, product demands and task performance measures in the context of a capability-demand model of interaction.

7.2 Study Design

The investigations presented in the previous chapters highlighted the interacting, complex nature of capability loss in the disabled population. The interaction of these capability losses ultimately determines the difficulties experienced in using products in daily life. Most studies in the literature examine certain classes of capability loss in isolation (Davis, 1994; Evamy & Roberts, 2004; Kanis, 1993; Langdon et al., 2007). However, the issue of multiple capability losses has not been adequately studied. In addition, most studies tend to focus on a single product family for evaluation at a time (Cardoso, 2005; Langdon, Lewis, & Clarkson, 2010; Waller, Langdon, & Clarkson, 2009b; Yoxall et al., 2010b). In reality, people use one product from different product families to build the product ecology found in the home environment.

A different approach was taken in the design of this study where both issues of *multiple capability loss* across *multiple products* were investigated. This novel multiple-capability, multiple-product study design was implemented to explore the range of relationships emerging from interaction. This is in contrast to studies that aim to identify the best product in a product category e.g. the best toaster of a series of toasters. In addition, another novel feature of the study was the investigation of the concept of difficulty as it relates to multiple capability loss and product usage. Recognising the importance of subjective factors in the context of use (Yoxall et al., 2010c), the perceived level of difficulty in performing various actions with the product was measured and related to user capabilities and product demands.

The overall aim of the study was to investigate the relationships between user capabilities, product demands and task performance measures in the context of a capability-demand interaction model. The design of the study is illustrated visually in Figure 7-1. The right of the diagram (*measurement/model context*) illustrates that firstly measures of users' sensory, cognitive and motor capabilities were recorded prior to performing tasks with four consumer products. The study design aimed to characterise users by measuring the *multivariate capability profile*, or performance envelope, of each user. Product demands were also recorded prior to users performing tasks, for example sizes and contrast of text and interface features, forces and motions required to operate the product, and the number of steps required to perform the task. Thus with both user capability and product demand data collected, the level of demand placed on each user was calculated.

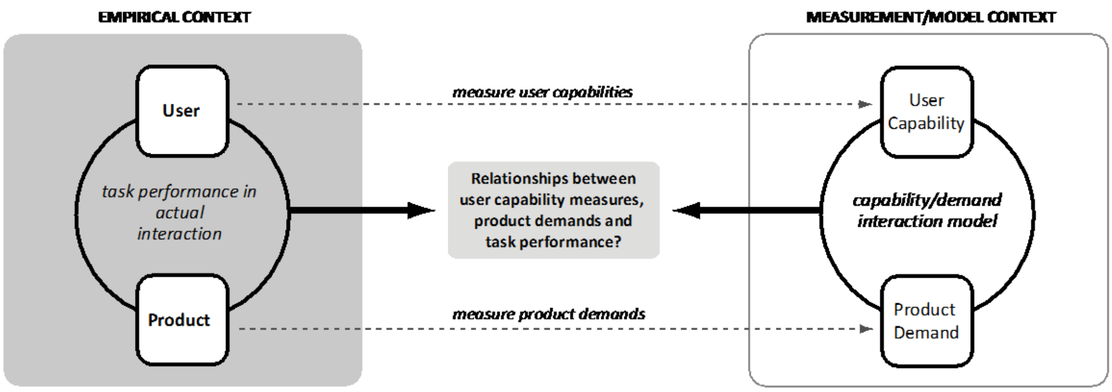


Figure 7-1 Study design in diagrammatic form relating the empirical context to a measurement context

The left of the diagram (*empirical context*) illustrates that outcome measures such as time taken, errors, and difficulty ratings were derived when users actually used the various products chosen for the study i.e. after actual interaction. These measures were compared to the measures of capability to investigate what relationships exist with an eye to developing a suitable model for inclusive interaction. Hence the relationships between task performance measures and user capabilities at set levels of demand were visualised. The investigation also shed light on the interaction of various user capabilities in task performance and the implications of this for using capability data for design evaluation.

7.2.1 Study Planning

Based on the logic of the study design, the necessary pre-work such as selecting the study location, sourcing participants, selecting products and tasks and sourcing testing equipment was carried out. An application was made for ethical approval from the Cambridge Psychology Ethics Committee and approval for the study was granted (see Appendix 3: Ethical Approval Letter). The usability laboratory on the 1st Floor of the William Gates Building (Computer Science) at the University of Cambridge was identified as an ideal location for the study given that the building is accessible with disabled parking spaces near the facility entrance (Figure 7-2).

The lab consisted of two rooms, one of which was an observation room and the other an open plan space that could be configured based on need. The open plan space was used for the study. In addition, there was a communal area just outside the lab which was used for welcoming, interviewing and debriefing participants. A café was located downstairs which was used for refreshments during the study break.

Figure 7-2 View from outside the usability lab on the left and from inside the lab on the right.

7.2.2 Selecting Products and Tasks

Four products were chosen for the study based on the criteria that they should (1) represent a wide range of product interfaces, (2) utilise common controls familiar to users (i.e. products that they were likely to have owned or used before) and (3) typically reflect products used in activities of daily living. A clock radio, a mobile phone, a blender and a vacuum cleaner were selected (Figure 7-3). Table 7-1gives more information on the interface elements for each product and the environment where each is commonly found. The clock radio was selected to characterise similar products such as CD players and alarm clocks. The mobile was selected to characterise similar products such as remote controls and other hand held devices. The blender was selected to characterise similar products with handles found in the kitchen like jugs, kettles, and coffee makers. The vacuum cleaner was selected to characterise similar products such as brooms, mops, lawn mowers and even dustbins that have to be pushed and pulled around. It was felt that these four products characterised a good range of interface features, product categories and demanded actions for use in the study.

Figure 7-3 Products selected for the study (from the left): Matsui Clock Radio, Siemens Mobile Phone, Breville Blender, and Panasonic Vacuum Cleaner.

Table 7-1 Products and tasks for experimental study

	Clock radio	Mobile phone	Blender	Vacuum cleaner
Activity	Personal management	Communication	Cooking/Feeding	Cleaning
Type of Product	Transportable	Handheld	Transportable	Transportable
Description/Interface Elements	LCD Display, small buttons, small switch, small dial, small chassis	LCD Screen, small buttons, small (handheld) chassis, menu system	Large rotary control, jug/carafe, midsized cassis, handle	Midsized to large chassis, large handle, large buttons (hand and foot), slider switch
Environment	Bedroom, bathroom, study, kitchen	On the person, multiple environments including: bedroom, office, car, outdoors	Kitchen (domestic and industrial)	Domestic and industrial environments with carpeted and hard flooring
User Task	Setting the time to 4.30 PM	Taking the ringer off via the menu	Making a smoothie with a banana and water	Vacuuming a piece of carpet till clean

It was also felt that these choices would control the effects of learning to some degree. In addition, it was hoped that this strategy would reduce ceiling and floor effects, especially when dealing with disabled participants with a range of capability. After four products were selected, a high level task analysis was performed on each product with the aim to select one task for each product that was required for daily use. The selected tasks are shown in Table 7-1. For the clock radio, blender and the vacuum cleaner, the task selected reflected the main function of the product. For the mobile phone, it was felt that the ringer task would be the best choice to investigate the challenge of menu browsing and selection. Descriptions of each task are given in Appendix 7: Task Sequences.

From the task analysis, a list of the main actions that were necessary to successfully complete the task was used to develop the capability tests that would be administered to participants. The list of actions was also used to develop a list of questions on experienced difficulty which users could rate after performing the task. Though it was anticipated that users would not perform each task in exactly the same way (for example employing different coping strategies), there are certain actions that are essential in achieving the goal, and these actions were selected to appear on the list.

7.2.3 Measuring Product Demands

For the list of essential actions for each product mentioned in the previous section, product demand parameters were measured. These measures included the dimensions of physical features, colours and contrast values, and the forces required for operation. See Appendix 8: Product Demand Values for a complete list of all recorded measurements. Table 7-2 lists the measuring devices used to extract measurements from the four products.

Table 7-2 Measures and Measurement Devices for Product Demands

Measure	Measuring Device
Dimensions	Ruler, Measuring Tape, Vernier Calliper
Contrast	Contrast Cards (see Appendix 6: Contrast Cards)
Forces	Mecmesin Digital Advanced Force Gauge (AFG 500N)

7.2.4 Measuring User Capabilities

Capability tests were specified and obtained to match the sensory, cognitive and motor demands of the essential actions. These tests were researched and purchased from various companies in the UK and the USA. Pictures of the testing equipment are shown in Figure 7-4 and Figure 7-5.

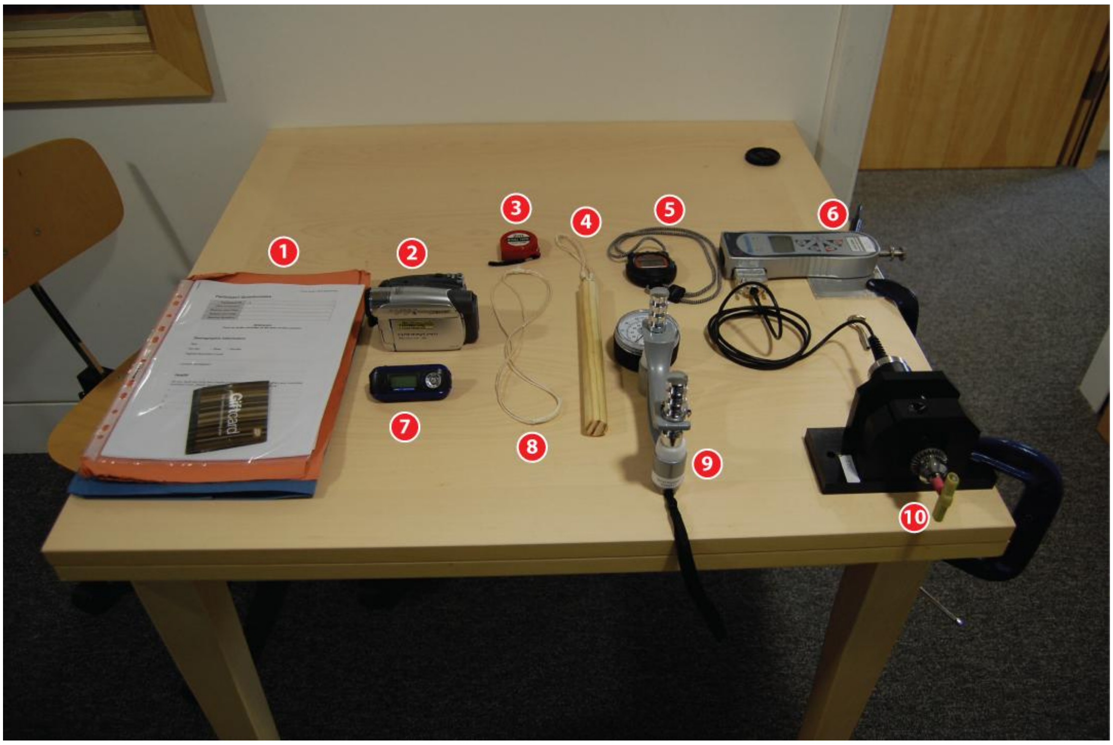


Figure 7-4 Study equipment: (1) Study pack with gift card, consent form and data collection sheets, (2) Sony Video Camcorder, (3) Measuring Tape, (4) Wooden handle, (5) Stopwatch, (6) Mecmesin Digital Advanced Force Gauge (AFG 500N) with clamp, (7) MP3 Audio Recorder, (8) String for measuring finger pull forces, (9) Jamar Grip Strength Measurement Device, (10) Torque measurement device with clamp.

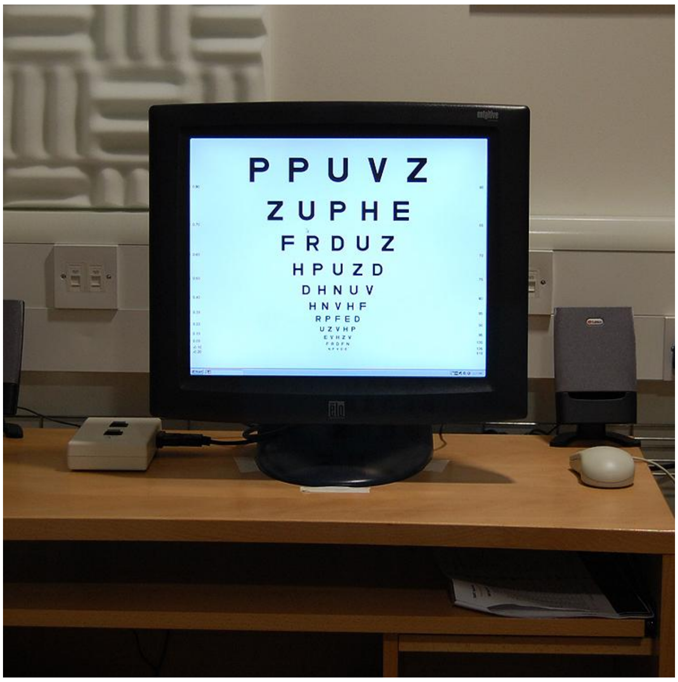


Figure 7-5 Computer based vision test on a touch screen monitor

A list of the testing equipment used for individual capability tests is given in Table 7-3.

Table 7-3 Capability Tests and Testing Equipment

Capability Test	Measures and Testing Equipment
Vision: Visual Acuity	Test Chart Pro Software
Vision: Contrast Sensitivity	Test Chart Pro Software
Hearing: Hearing Thresholds	NCH Tone Generator
Cognition: Verbal Working Memory	Digit Span Paper based test
Cognition: Spatial Working Memory	CANTAB Beclipse Software
Cognition: LTM	CANTAB Beclipse Software
Cognition: Reaction Time	CANTAB Beclipse Software
Motor: Grip Strength	Jamar Hydraulic Hand Dynamometer
Motor: Pinch Strength	Jamar Pinch Gauge
Motor: Finger forces, Push/Pull Forces and rotational forces	Mecmesin Digital Advanced Force Gauge (AFG 500N) with clamp (accuracy ±0.1% of full-scale)

The Mecmesin Digital Advanced Force Gauge (AFG 500N) with torque sensor came with a calibration certificate from the manufacturer. The calibration of the device was further checked in the workshop at the Cambridge Engineering Department against a variety of known weights across the intended force range of measurement in the study. This was also done for the grip strength and pinch meter. Data collection forms were designed to record the capability data as the tests were performed (see Appendix 10: Experiment Data Collection Sheets). The following sections detail the tests, procedures and measures for sensory, cognitive and motor capability testing.

7.2.4.1 Measuring Sensory Capabilities

Visual capability was assessed via the measurement of two visual functions i.e. distance visual acuity (using a letter chart) and contrast sensitivity (using letters of various contrasts) as shown in Figure 7-6. Visual acuity scores were measured using a logMAR (log Minimum Angle of Resolution) chart and the Visual Acuity Rating (VAR). This VAR is advantageous in that it provides a simple method for scoring patients using the logMAR chart. The VAR score can be converted to a logMAR value, which in turn relates to a high contrast letter size.

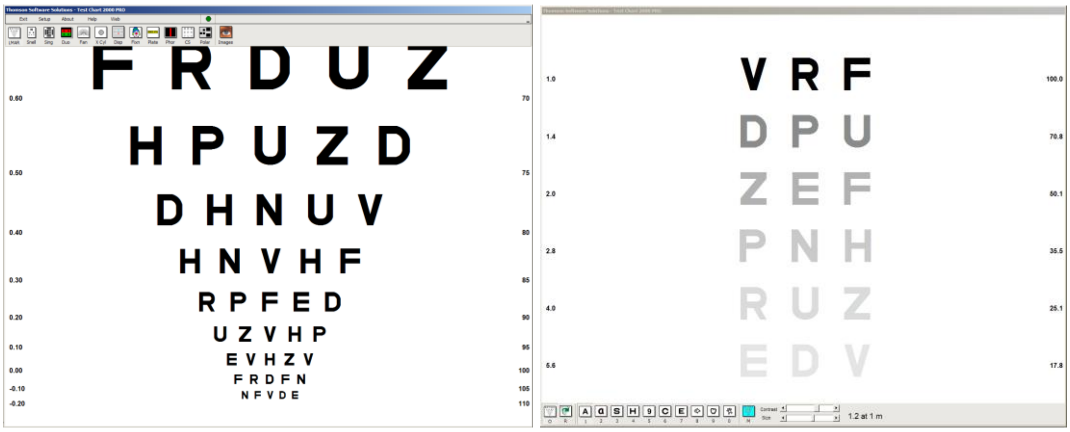


Figure 7-6 Screenshots of visual acuity test using a logMAR chart on the left, and a contrast sensitivity test using a test chart of letters with varying sizes and contrasts on the right (Test Chart Pro Software).

Participants were seated 3 metres from the screen. The Test Chart 2000 Pro manual describes the scoring procedure as follows: "The 0.00 (6/6) row is given a score of 100 and each letter has a score of 1. For example if a patient reads all the letters down to the 100 row but none on the row below, their score is 100. If they get one letter wrong on this row, their score is 99, two letters wrong 98, three 97, four 96. If no letters are read on this row, their score is 95. If they read all of the letters on the 100 row and one letter on the row below, their score is 101, two letters on the line below 102 etc. This is not only easier to score, but also produces a number that will have some meaning to the patient – 100 being normal, 105 being better than average, 95 being slightly below average etc." Thus a threshold VAR score and a comfort VAR score were obtained for each participant, provided that they could perform the test.

The contrast sensitivity test shown on the right of Figure 7-6 required participants to read a randomised set of letters at six different sizes (logMAR 0.4, 0.7, 1.0. 1.3, 1.6 and 1.8) while seated at 1 metre from the screen. Letters were displayed in triplets of decreasing contrast from the top to the bottom of the screen. Participants were asked to read the letters from the top to the bottom until they could no longer read two out of the three letters displayed. The contrast of this row was recorded as the minimum contrast the participant could see at that size.

Rough measures of hearing capabilities were taken using NCH Tone Generator software, where white noise was generated at 73dB, 67dB, and 57dB (measured with a sound level meter at the ear), and participants were asked if they could hear the noise at successively quieter levels. All participants could hear the noise at 57dB, and since measured noise values from the products were all above 73dB, it was felt that participants would have no difficulty with hearing tasks.

7.2.4.2 Measuring Cognitive Capabilities

Cognitive capabilities were measured using 4 different tests. The first of these, verbal working memory, was assessed using a standard digit span test, where participants were asked to repeat a string of random digits of increasing length till an error was made in the repetition of the sequence. The length of the longest correct sequence recalled was recorded as the digit span. The next three tests utilised CANTABeclipse software running on a desktop computer with a touch screen monitor. CANTABeclipse is computer based cognitive testing software from Cambridge Cognition Ltd (www.cantabeclipse.co.uk) which is widely used in cognitive and clinical research. Table 7-4 lists the 4 CANTABeclipse tests used.

Table 7-4 Description of CANTABeclipse cognitive tests

CANTABeclipse Test	Description
Motor Screening (MOT)	Screens for visual, movement and comprehension difficulties
Spatial Span (SSP)	Assesses working memory capacity
Reaction Time (RTI)	Measures speed of response
Graded Naming Test (GNT)	Gives a measure of semantic memory by assessing object-naming memory

A motor screening test (MOT) was first run to ensure that participants possessed sufficient capacity to perform the three tests that follow by screening for visual, movement and comprehension difficulties. It is administered at the beginning of a test battery and serves as an introduction to the touch screen. If a subject is unable to successfully perform the test, it would be unlikely that they will be able to complete other tests successfully. In this test, participants touched a flashing cross which was shown in different locations on the screen.

Visuo-spatial working memory was assessed using the Spatial Span (SSP) test. Spatial Span assesses working memory capacity, and can be considered a visuo-spatial analogue of the digit span test. It is a computerised version of the Corsi Blocks task and gives a measure of frontal lobe functioning. The test procedure is as follows from the software manual: “White squares are shown, some of which briefly change colour in a variable sequence. The subject must then touch the boxes which changed colour in the same order that they were displayed by the computer. The number of boxes increases from 2 at the start of the test to 9 at the end, and the sequence and colour are varied through the test.” Span length (the longest sequence successfully recalled), was used as the outcome measure for this test. Figure 7-7 shows a screenshot of this test.

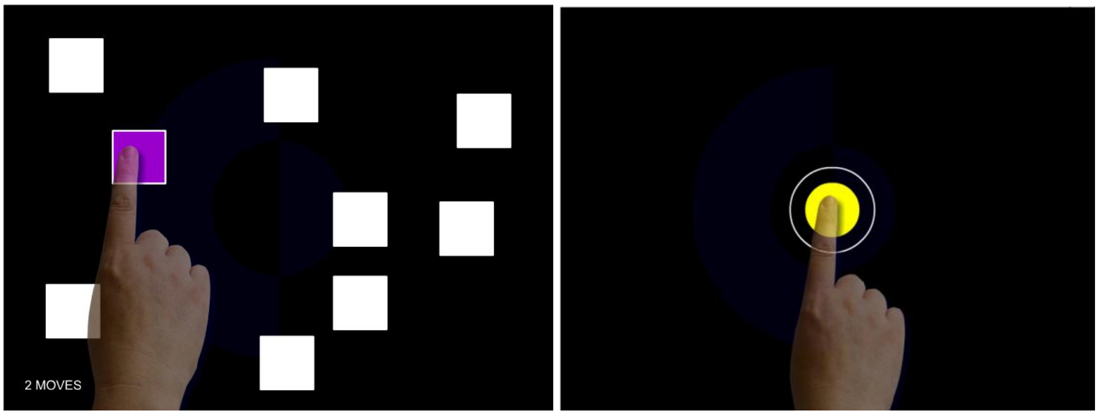


Figure 7-7 Screenshots of the Spatial Span Test (SSP) on the left and the Reaction Time Test (RTI) on the right. Interactive simulations and descriptions of CANTABeclipse tests are available at www.cantabeclipse.co.uk.

Reaction time (RTI) is a latency test designed to measure speed of response to a visual target. It was assessed using a task comprising five stages, which requires increasingly complex chains of responses. This test used a button on a press pad connected to the computer in addition to the touch screen functionality. The software manual lists five stages of the test:

"In the first stage, the subject has to touch the screen when a yellow dot appears in the centre of the screen, neither touching too soon nor too late (5 out of 6 correct for success). In the second stage, the yellow spot may now appear in any of five locations (5 out of 6 correct for success). In the third stage, the subject is required to hold down the press pad button until the yellow spot appears in the centre of the screen, and then must touch the screen where the spot appears. In the fifth and final stage, the choice reaction task is again introduced, and by this stage the subject has been trained to hold down the press pad button until the spot appears, then release the press pad button and touch the position on the screen where the spot was presented (5 out of 6 correct for success). If the subject fails to reach the criterion on any of these stages except the first stage, the task terminates."

The first three stages are for practice only, and stages four and five are the test trials. The outcome measure used for the RTI was five-choice reaction time. Figure 7-7 shows a screenshot of this test.

Long term memory was assessed using the Graded Naming Test (GNT) which is used extensively in cognitive neuropsychology. This test assesses object-naming ability, and is also graded in difficulty to allow for individual differences. The test consists of 30 different line drawings that are displayed on the screen, one at a time. The participant must identify the object depicted in each drawing. The total number of correct responses (as a percentage) was used as the outcome measure. Note that this test was culturally biased to a UK population at the time of testing. For all tests, the test scripts provided in the software manual were used to administer the test. Test scores were automatically logged in a database within CANTABeclipse, and these were retrieved at the end of the testing session. Figure 7-8 shows graphical output from the software for each participant, giving raw scores, standard scores and graphical comparisons with normative population data.

Subject ID	02	Gender	Male
Age	N/A	User name	Administrator

Test	Measure	Raw score	Standard score	Standard score chart	Better than	Good as or better than	Population diagram	Comparison basis			
								Age	NART	M/F	N
MOT	*Mode* clinical *Test start time* Jul 2, 2007 2:36:08 PM *Test duration* 1 min 9 s										
	Mean error	8.21	0.36		95-100%	95-100%		All	All	M	237
SSP	*Mode* clinical *Test start time* Jul 2, 2007 2:37:17 PM *Test duration* 9 min 6 s										
	Span length	5	-0.56		15-20%	45-50%		All	All	M	280
RTI	*Mode* clinical *Test start time* Jul 2, 2007 2:46:24 PM *Test duration* 5 min 52 s										
	Five-choice reaction time	307.25	0.75		75-80%	75-80%		All	All	M	102
GNT	*Mode* clinical *Test start time* Jul 2, 2007 2:52:17 PM *Test duration* 4 min 1 s										
	Total correct	28	1.61		95-100%	95-100%		All	All	All	400

Figure 7-8 Graphical results from CANTABeclipse for one participant showing raw scores and comparisons with normative data

7.2.4.3 Measuring Motor Capabilities

A pictorial example of a user performing the motor capability test battery is given in Figure 7-9. In addition to threshold or liminal values (maximum values), comfort levels were also recorded for motor capabilities.

Figure 7-9 User performing motor capability tests: (01, 02) Twisting force measurement, (03) Grip strength measurement, (04, 05) Finger push force measurement, (06, 07) Pull and push force measurement, (08) Finger pull up force measurement and (09) Grasp and pull up force measurement.

For example, users were first asked to push with their finger as hard as they could for the finger force test. Then they were asked to push up to the point where they felt comfortable or felt the onset of discomfort. Tests were repeated twice and the average value taken. The figure illustrates the postures and orientations for the different tests. These were designed to closely match, where possible, the actions required when using the four consumer products.

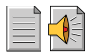
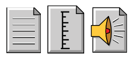
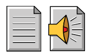
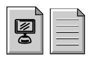
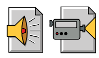
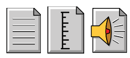
7.2.5 Sampling Users

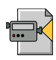
Users were theoretically sampled while trying to cover a range of ages and capabilities. Participants were contacted through three main organisations/charities: University of the Third Age in Cambridge (U3A), CAMSIGHT and the Hester Adrian Centre (Papworth Trust). Sites were visited and flyers were distributed and posted at various locations. For participating, participants were renumerated with a £10 Boots voucher and their transportation costs were covered. Participants were also provided with refreshments during the study. A consent form was developed for the study which was approved by the Cambridge Ethics Committee (see Appendix 4: Study Consent Form). 19 participants in total were recruited.

7.2.6 Study Procedure

The study procedure for each participant is illustrated in Table 7-5. The table also shows the type of data collected at each stage.

Table 7-5 Outline of the study procedure

	Stage	Data Capture	Data Format
1	Welcome, Introduction, Consent and Remuneration	Participants signed consent form Participants given Boots Voucher	
2	Background Questionnaire	Questionnaire	
3	Experience Ratings and Questionnaire	Participants Rated Product Experience	
4	Mental Models for Each Product	Verbal Protocol; Audio Recording	
5	Sensory, Cognitive and Motor Capability measurement	Computer Testing Written	
6	Performance of one task with each of four products (randomised)	Video Capture	
7	Difficulty Ratings for actions after each task	Visual Analogue Rating Scale	
8	Debriefing	Problems and Questions?	

KEY:  Textual  Rating Scale  Computer  Audio  Video

Pilot runs were conducted with 3 users prior to the commencement of the actual study in order to practise the procedure and resolve unexpected problems with the instructions, flow, and measurement equipment. The different stages are outlined in more detail below:

Participants first signed a consent form (large print forms were available). They were then asked background questions to gather demographic, medical and product experience information. Participants were also asked to rate their experience with four consumer products and to describe how they would go about using these products to perform tasks (one task per product). Data was recorded on a questionnaire sheet and backed up with an audio recorder.

Secondly, a battery of capability tests was administered. These tests included sensory tests (visual acuity, contrast sensitivity, hearing level), cognitive tests (verbal and spatial working memory, speed of processing, long term memory) and motor tests (force exertion in various positions, balance and walking speed). It was emphasised to the participants that they were not being tested; rather information was being collected to investigate how well the products matched their capability (Boren & Ramey, 2000). Participants had a short break after the sensory and cognitive capability assessments were performed. Some of the participants took other breaks during the capability testing session if they became tired, while other participants chose to continue straight through without a break till all of the capability tests were completed. All capability testing data was recorded on a pre-designed testing sheet. Data from the computerised cognitive capability tests was stored in a computer database and later exported for analysis.

Finally, participants were asked to perform one task each with four consumer products: a clock-radio, a mobile phone, a blender, and a vacuum cleaner. The four tasks were randomised for each participant. At the start of the tasks, participants were informed that they could stop at any time for any reason. While tasks were performed, their performance was videotaped. On completion of each of the four tasks, participants rated the level of difficulty and frustration experienced for selected actions on the task with the aid of a graphic difficulty scale.

A visual rating scale was developed and used to help participants to quantify their difficulty with various actions in using the products (see Appendix 5: Difficulty Scale). The scale ranged from 0 to 100 and utilised graphic smiley faces to clarify the scale end points. This facilitated ease of rating difficulty, especially for participants with reduced cognitive capability. Care was taken to ensure that participants were not asked leading questions. For example, participants were asked neutral questions such as how ‘easy or difficult’ it was to perform various actions (Boren & Ramey, 2000). A study checklist was used to guide the researcher through the study procedure (see Appendix 9: Experimental Study Checklist). In the next section, the data analysis procedures will be discussed before the analysis of the data is presented.

7.3 Results

7.3.1 Data Analysis Procedure

The study data was entered into SPSS 17 for further analysis (Figure 7-10). This included the demographic and questionnaire data (including ratings of product experience) collected from interviews, and the sensory, cognitive and motor capability measures. In addition, the ratings of difficulty and frustration for each product task were also entered for each participant.

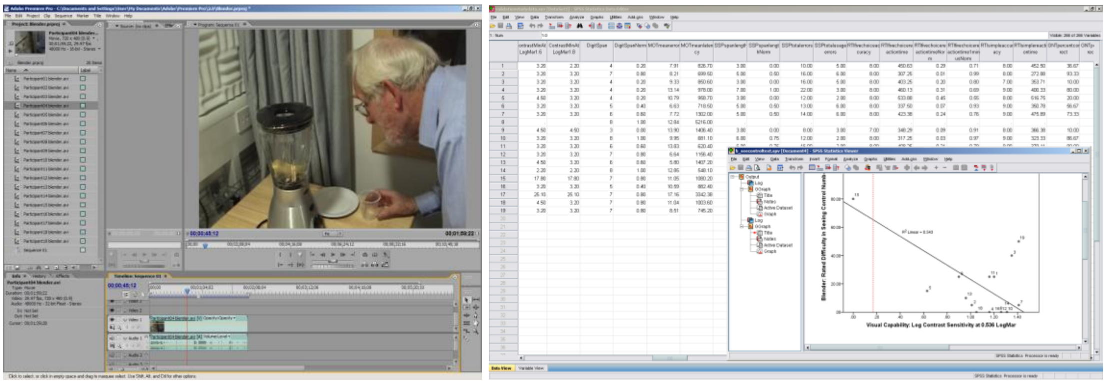


Figure 7-10 Data being analysed in Adobe Premiere Pro 2.0 and SPSS 17

All video data of the product interactions was transferred from miniDV tape to computer video files (AVI) for analysis using Adobe Premiere Pro 2.0 video editing software (Figure 7-10). This resulted in 75 video files that were subsequently analysed (19 participants x 4 videos each = 76 videos, with one participant not attempting a task resulting in 75 videos). Each video was analysed and task outcome measures were extracted. These included task success/failure, task times, and errors made. Errors were classified using an error taxonomy developed for the SHERPA: Systematic Human Error Reduction and Prediction Approach (Stanton & Baber, 2002). Each video was firstly reviewed in its entirety, and then errors were noted on a second viewing. The video was viewed a third time to ensure that no errors had been missed. A detailed classification of the errors was not carried out as it was felt that it was not necessary in light of the aims of the study. All task outcome measures were finally added to the SPSS data file giving a complete data set of 195 measured variables for 19 cases. After calculating other derived variables (for e.g. normalised scores, capability-demand values), the final data file consisted of 266 variables.

All data collected was stored and treated following the University of Cambridge's research data policies in accordance with the Data Protection Act 1998 (the eight data protection principles). This guidance was followed in collecting data (consent obtained for all uses of the data), storing the data (anonymised for data files and stored in a secure location both physically and electronically) and in using and presenting the data. In the following sections, an analysis of the data will be presented. Firstly, an overview of participant demographics and measured capability distributions will be given. Secondly, the results of the task outcomes will be described (task success/failure, difficulty scores, times and errors). Finally, an analysis of capability-demand relationships will be presented via correlation scatter plots of rated difficulty versus measured capability for constituent actions.

7.3.2 Participant Characteristics

Figure 7-11shows all 19 participants in the study performing the blender task. Table 7-6 gives the educational, occupational and medical background of each participant.

Figure 7-11 Study participants 1 to 19 using the blender

Table 7-6 Educational, occupational and medical background of the 19 study participants

#	Age	Sex	Education	Occupation	Medical conditions and aids
1	62	F	Secondary School	Hester Adrian	Thyroid medication, glasses
2	64	M	Degree	Retired School Teacher	Early Parkinsons, glasses
3	50	M	Training school	Hester Adrian and Tesco	Epillepsy
4	82	M	Degree	Retired Engineer	Glasses
5	65	F		Hester Adrian	Tinnitus, balance problems
6	71	F	Diploma	Retired bookseller	Back problems (age related), high blood pressure
7	44	M	GCSE	Hester Adrian	Spinal Bifida, glasses
8	47	F	Degree	Local government officer	High blood pressure, macular degeneration, guide dog, cane
9	54	F	School for learning disabilities	Hester Adrian	Glasses
10	61	F	Degree	Social Research Interviewer (NATCEN)	RSI in right hand, short sightedness, dupitrons in hands, glasses
11	63	M	Degree	Retired technology consultant	-
12	62	F	Degree	Retired Secretary	Osteoporosis, glasses
13	66	M	A'Levels	Retired Army Officer	Glasses
14	72	F	Teacher's Diploma	Retired primary school teacher	Genetic condition, tinnitus, lumps on hands, glasses
15	74	F	O'Levels	Self employed	Genetic macular degeneration, magnifying glass, CCTV, text to speech
16	61	F	Postgraduate certificate of education	Retired teacher	-
17	65	M	Degree	Retired Architect and town planner	Diabetes, thyroid deficiency, retinitis pigmentosa, magnifying glass, text to speech
18	64	F	Degree, Diploma in social work	Retired Probation Officer	Chronic Fatigue Syndrome, minor stroke, arthritis in spine and wrists, glasses
19	64	F	GCSE	Retired medical secretary	Arthritis in hands, structural deformities in feet, glasses

Figure 7-12 shows the age and sex distribution for the study sample (Mean Age=62.68, SD=9.20) ranging from 44 years to 82 years. Of the 19 participants that participated in the study, 12 were female and 7 were male.

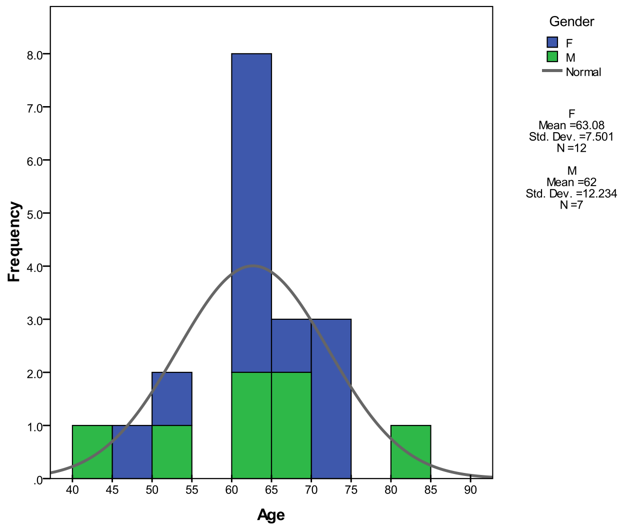


Figure 7-12 Age and sex distribution of participants

Figure 7-13 shows the educational background of the study sample. 3 participants attended special schools for disability training, 5 participants had a secondary school education, 2 participants studied to the diploma level, 8 participants had a university degree and one participant had a postgraduate qualification. 5 of the 19 participants worked at the Papworth Trust's Hester Adrian centre which provides job opportunities for people with disabilities and illnesses. 11 participants were retired from their jobs which included teaching, engineering, business, government services, military services, and computer technology. 3 participants were still employed in local government, social research interviewing, and self-owned business. Thus participants presented with a wide range of backgrounds which was desirable in the study.

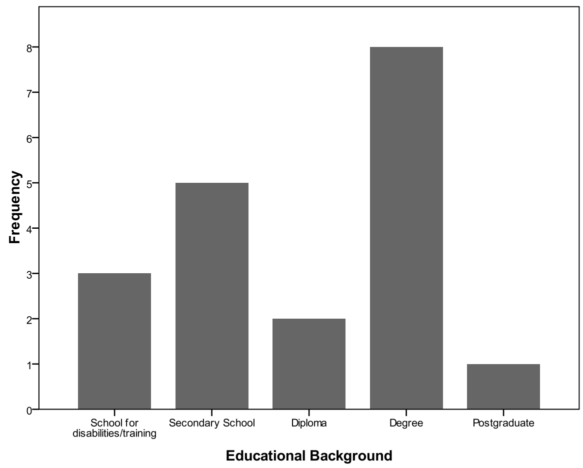


Figure 7-13 Educational background of study participants

When asked about current medical conditions, participants reported a range of conditions (including some associated with ageing): tinnitus, macular degeneration, retinitis pigmentosa, parkinson's, epilepsy, arthritis, osteoporosis, structural deformities, spinal bifida, high blood pressure, RSI, diabetes, thyroid deficiency, chronic fatigue syndrome and minor stroke. Participants were also asked to report if they had any problems with their general vision, hearing, memory, hand use and walking. 6 of 19 participants (31.6%) reported problems with their vision, 5 (26.3%) participants reported problems with their hearing, 6 (31.6%) participants reported problems with their memory, 6 (31.6%) participants reported problems with using their hands, and 6 (31.6%) participants reported difficulties with walking. In addition, participants used various aids to assist with daily functions. One participant used a hearing aid, 11 participants used glasses, two participants used magnifying glasses and text-to-speech systems, and one participant used a guide dog with a cane. For movement, one participant used walking sticks and another used a wheelchair/walker.

7.3.2.1 Sensory Characteristics

Figure 7-14 shows the distribution of VAR scores in the study sample. The graph on the left of the figure shows that the distance acuity (M=88.7, SD=17.3, N=18) of most participants ranged between 80 and 100. Two participants had significantly lower acuity scores, and one participant could not read the chart at all. The graph on the right of the figure shows the distribution of "comfort acuity" scores (M=78.5, SD=19.2, N=17). This was measured by asking participants to pick a line on the logMAR chart that they felt comfortable reading without strain. The graph shows lower VAR scores for comfort acuity, which translates to larger letters. Most of the participants had good to fair visual acuity scores given the age group and vision correction devices used (mostly glasses). However, two participants had particularly poor vision with VAR scores of 55 and 35, and one participant's vision was so low that she could not perform any of the vision tests.

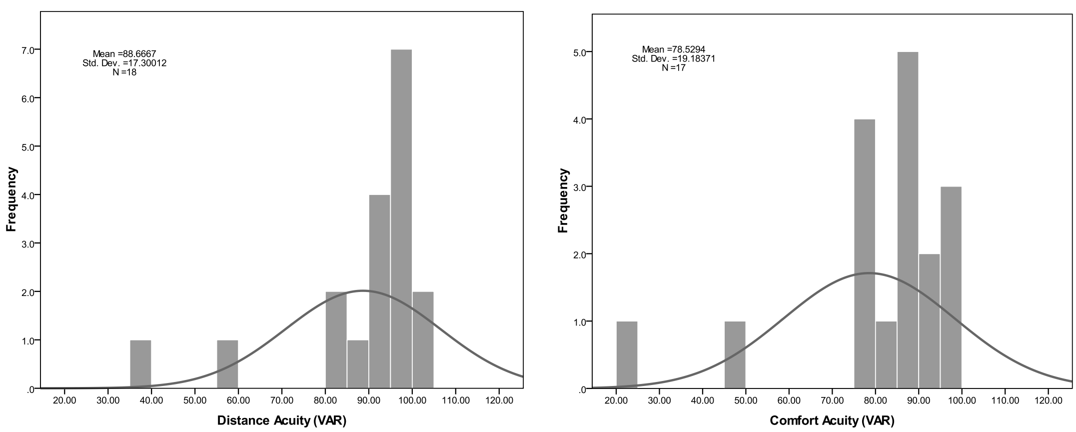


Figure 7-14 Histograms of distance visual acuity on the left and a comfort acuity on the right

Figure 7-15 shows the results of contrast sensitivity testing at different sizes of letters. The graph plots mean minimum contrast levels in the sample at 6 different letter sizes (error bars show standard error of the mean). As expected, at larger text sizes, participants could recognise letters of low contrast, e.g. 5.2% at 1.8 logMAR on average. At smaller text sizes, the average contrast threshold is raised e.g. 20.2% at 0.4 logMAR. The variability in the minimum contrast threshold also increases as the letter size decreases.

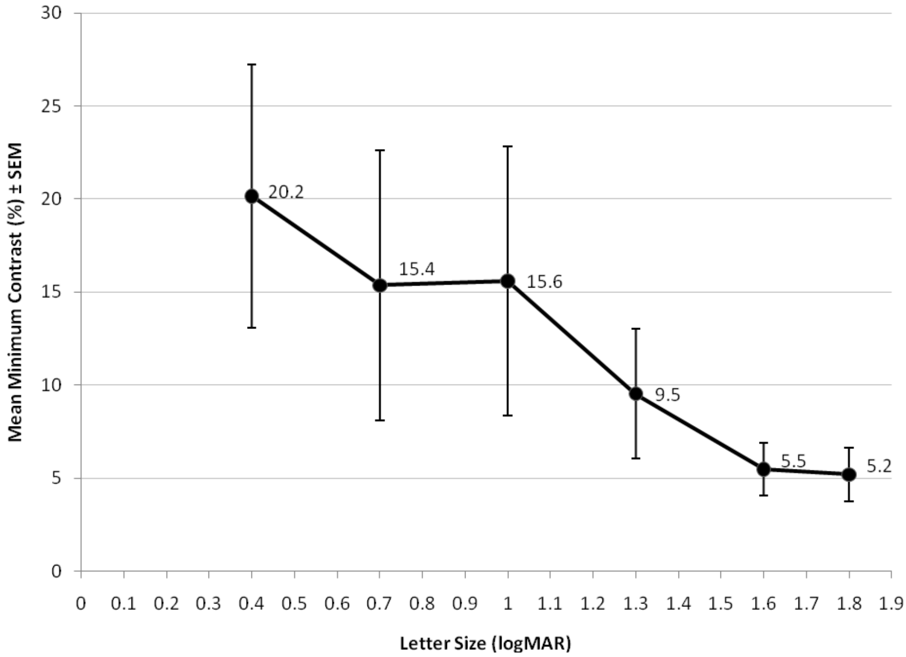


Figure 7-15 Mean minimum contrast levels (%) at different sizes of text (logMAR) for the study sample.

7.3.2.2 Cognitive Characteristics

Figure 7-16 shows the distribution of the 4 cognitive capability measures. The distribution of digit span scores (M=5.95, SD=1.58, N=19) shows that a score of 7 had the highest frequency, while a score of 4 had the second highest frequency. 5 participants had fairly low digit span scores of 4 or less. Two participants could not successfully perform the MOT test (and thus the three subsequent cognitive tests) because they could not see the objects on the screen. Spatial span scores (M=4.8, SD=1.29, N=17) show that scores of 5 and 6 were the most frequent, followed by relatively low scores of 3. The histogram of five choice reaction time (M=412.2, SD=122.31, N=17) shows times between 400 – 450 milliseconds to be the most frequent in the sample. One participant had a particularly long reaction time above 800 milliseconds. This participant had poor vision which possibly accounted for this high score. The histogram for the graded naming task (M=62.5, SD=32.39, N=17) shows a large distribution of scores across the 100% range. Scores between 80-90% were the most frequent.

The cognitive capability scores indicate a reasonable spread over their respective ranges. This is in keeping with the aims of the study to investigate performance of users across a range of capability profiles.

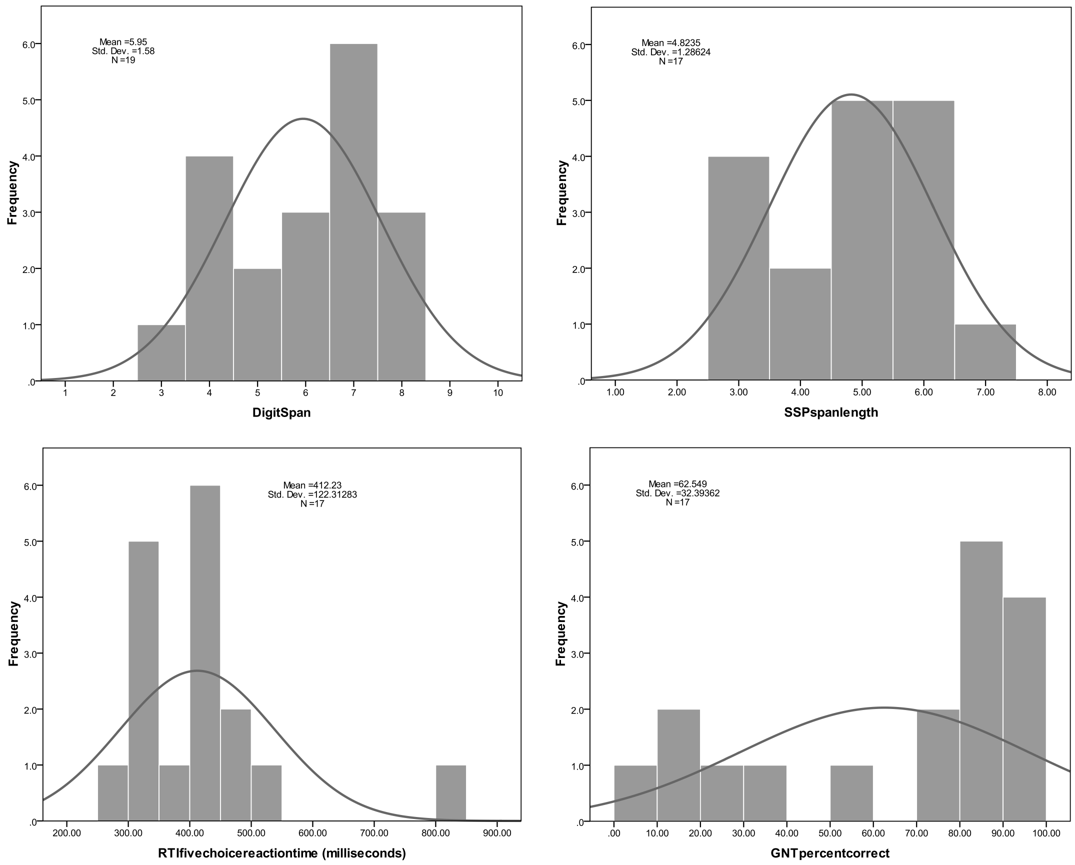


Figure 7-16 Histograms of cognitive variables measured: digit span, spatial span, five choice reaction time and graded naming task (GNT)

7.3.2.3 Motor Characteristics

Table 7-7 lists the descriptive statistics for the 37 motor capability variables recorded for each participant. All forces were measured in Newtons, rotational forces in Newton-metres and times in seconds. In addition to the forces listed, pinch strength between index finger and thumb for each hand was listed for measurement. However, the pinch gauge sourced for the study malfunctioned after the first session, and a replacement could not be sourced in time for the remainder of the study. Thus this capability measure was not captured. Since force measures were captured for each hand in different positions and grasps, it was possible to relate the motor actions captured on video, e.g. pushing, lifting etc., to the exact hand that was used. This led to accurate plots of relationships between motor capability and motor demand forces for each participant.

Table 7-7 Descriptive statistics for motor capability variables

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
RHGripstrengthComf	17	26.00	6.00	32.00	15.3529	7.58239
LHGripstrengthComf	16	36.00	4.00	40.00	16.8750	8.67852
RHGripstrengthMax	19	48.00	2.00	50.00	24.4211	11.54852
LHGripStrengthMax	18	44.00	4.00	48.00	25.2778	12.05285
RHRotateAnticlockwise Comf	16	3.518	.458	3.976	2.09350	1.017910
RHRotateAnticlockwise Max	18	3.436	1.262	4.698	2.71978	1.211432
RHRotateClockwiseComf	16	2.716	1.008	3.724	2.05388	.931927
RHRotateClockwiseMax	18	3.338	1.240	4.578	2.66556	1.048652
LHRotateAnticlockwise Comf	16	2.874	.812	3.686	1.84650	1.007383
LHRotateAnticlockwise Max	18	3.324	.708	4.032	2.30567	.964733
LHRotateClockwiseComf	16	3.254	.908	4.162	2.27850	.999058
LHRotateClockwiseMax	18	4.000	.986	4.986	2.92389	1.127125
RHFingerpushComf	16	72.50	12.40	84.90	40.0313	16.95998
LHFingerpushComf	16	83.00	10.70	93.70	41.4438	18.05676
RHFingerpushMax	19	94.80	9.30	104.10	54.7789	25.29351
LHFingerpushMax	19	71.30	15.20	86.50	46.3842	19.49907
RHThumbpushComf	14	138.20	11.90	150.10	63.0500	34.66183
LHThumbpushComf	14	118.10	21.20	139.30	65.4000	32.52400
RHThumbpushMax	17	137.50	10.00	147.50	78.8765	40.92092
LHThumbpushMax	17	123.20	15.30	138.50	77.4412	38.45093
RHPullComf	16	60.40	21.30	81.70	50.0312	18.53853
LHPullComf	16	64.30	22.40	86.70	56.5938	20.17455
RHPullMax	19	104.40	35.80	140.20	82.8316	29.54806
LHPullMax	19	95.50	45.70	141.20	92.0684	28.35591
RHPushComf	16	50.50	28.90	79.40	54.1188	14.49603
LHPushComf	16	61.60	27.60	89.20	51.5688	18.43673
RHPushMax	19	103.90	33.00	136.90	81.1053	30.48924
LHPushMax	19	104.80	23.90	128.70	74.2105	27.19299
RHLiftComf	16	153.00	23.60	176.60	78.8938	41.16161
LHLiftComf	16	190.30	28.90	219.20	89.9187	48.91558
RHLiftMax	18	178.00	48.30	226.30	120.1667	54.02398
LHLiftMax	18	231.90	33.00	264.90	121.5000	56.73900
FingerPullupComf	16	71.00	26.90	97.90	52.1563	23.48069
FingerPullupMax	18	128.70	4.40	133.10	72.2167	34.95963
BalanceTime	19	60.00	.00	60.00	15.6689	18.42498
WalkingSpeedComf	16	22.02	5.94	27.96	9.5563	5.53767
WalkingSpeedFast	14	8.63	2.25	10.88	5.6879	1.97470
Valid N (listwise)	10					

Figure 7-17 shows some illustrative motor capability graphs for grip strength, clockwise rotation, index finger push and pushing while grasping for each hand.

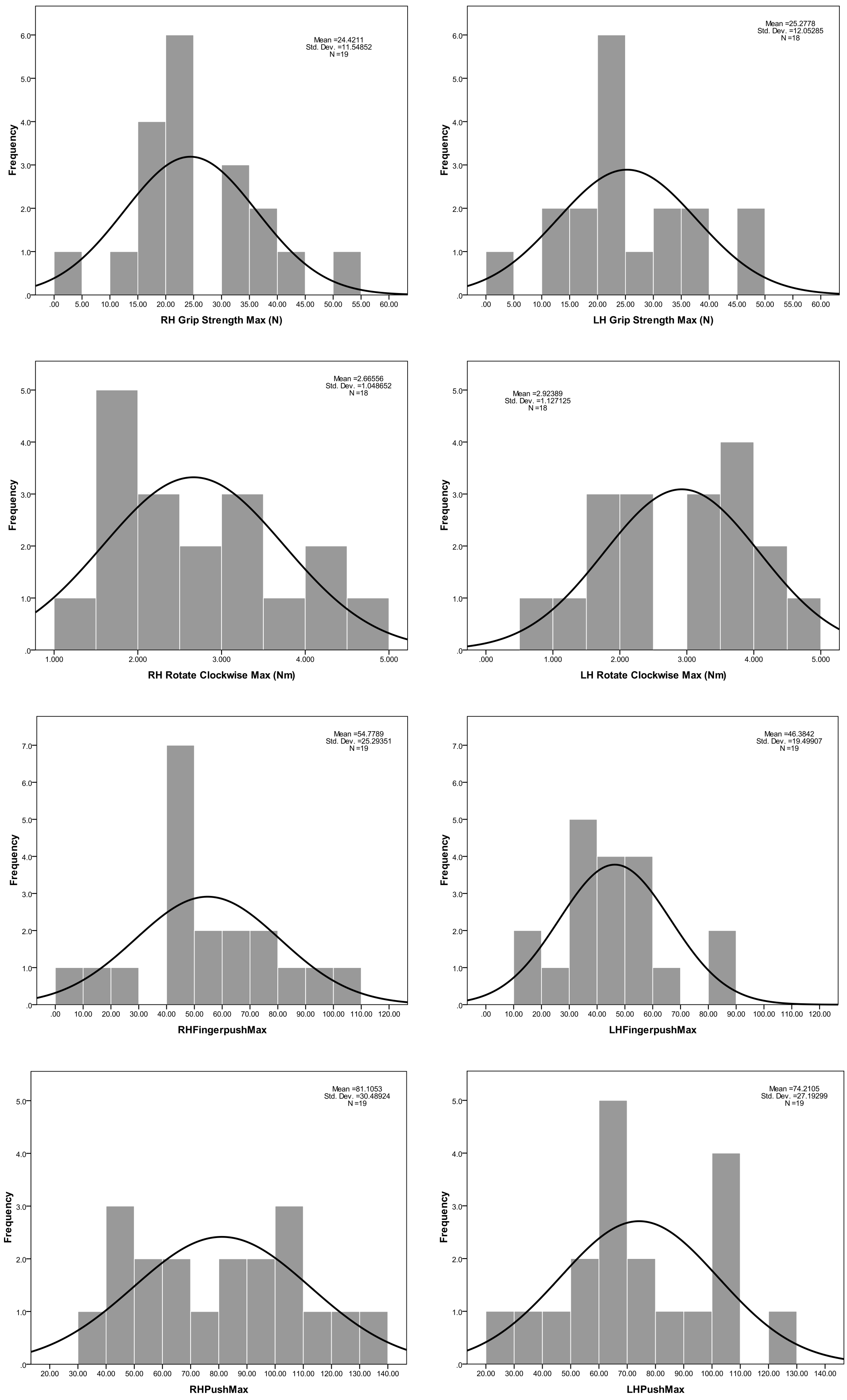


Figure 7-17 Motor capability histograms for grip strength, clockwise rotation, index finger push and push

The graphs show that the participants represented a good range of capabilities for each measure which was required in the study. In addition, there were observable differences in the distribution of each measure between the right hand and the left hand. This justified characterising each hand independently, especially given the unpredictable and asymmetrical effects of disease and trauma on hand/limb function.

7.3.3 Factor Analysis of Capability Variables

An exploratory factor analysis was performed on all sensory, cognitive and motor capability variables in order to ascertain the patterns of relationships among the data (Table 7-8). Only the maximum exertion variables were utilised (e.g. maximum push force) in this analysis. Factor loadings above or below ±0.4 were used for key variables in the components and Table 7-9 shows the total variance explained by the extracted components. Figure 7-18 shows a scree plot for the factor analysis. The point of inflexion is at 4 components, leaving a suggested 3 factor solution (Field, 2009).

Factor 1 explained 49.4% of the variance, comprising visual distance acuity, all motor force exertion variables and walking speed. This possibly indicates a general factor for *movement and force exertion* (including coordination for hand-eye operation and walking). Factor loadings for variables in the factor were very strong and explained almost half of the total variance. It is known that these capabilities tend to decrease with age.

Factor 2 explained 27.7% of the variance, comprising all the visual capability variables (acuity and contrast sensitivity), reaction time (as a measure of executive function) and long term memory. The loadings of right hand pull strength and finger pull up strength is just above the chosen threshold of above or below ±0.4. Therefore, this factor demonstrates a link between visual capability and cognitive processing involving long term memory retrieval. It is expected that if visual functions decline, the sensory processing will also decline as the quality of visual input is reduced. In essence, this factor represents the linked decline in sensory and general cognitive capability with disability and ageing.

Factor 3 explained 8.5% of the variance, and comprises three cognitive variables: digit span as a measure of working memory, span length as a measure of visuo-spatial working memory and long term memory. Though long term memory showed a loading on Factor 2, it has a higher factor loading on Factor 3 than Factor 2. Balance time and walking speed also have loadings above or below ±0.4 on this factor, though at smaller loadings than the three cognitive variables. Factor 3 therefore represents the decline in the memory component of cognition including working memory and long term memory. This decline also impacts upon the cognitive processing required for balance and walking, and is manifested in the large number of falls and locomotion problems among the elderly.

Table 7-8 Component matrix showing factor loadings for capability variables

Component Matrix[a]

	Component				
	1	2	3	4	5
DistanceAcuity	.515	-.851			
ContrastMinAtLogMar0.4		.937			
ContrastMinAtLogMar0.7		.952			
ContrastMinAtLogMar1.0		.944			
ContrastMinAtLogMar1.3		.953			
ContrastMinAtLogMar1.6		.941			
ContrastMinAtLogMar1.8		.929			
DigitSpan			.661		.677
SSPspanlength			.662	-.482	
RTIfivechoicereactiontime		.944			
GNTpercentcorrect		-.537	.740		
RHGripstrengthMax	.873				
LHGripStrengthMax	.927				
RHRotateAnticlockwise Max	.929				
RHRotateClockwiseMax	.867				
LHRotateAnticlockwise Max	.926				
LHRotateClockwiseMax	.979				
RHFingerpushMax	.806				.443
LHFingerpushMax	.771				
RHThumbpushMax	.858				
LHThumbpushMax	.907				
RHPullMax	.697	.448			
LHPullMax	.746			-.462	
RHPushMax	.899				
LHPushMax	.914				
RHLiftMax	.905				
LHLiftMax	.946				
FingerPullupMax	.837	.403			
BalanceTime			.451	.737	
WalkingSpeedFast	-.796		-.523		

Extraction Method: Principal Component Analysis.

a. 5 components extracted.

Table 7-9 Total variance explained by extracted components

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	14.831	49.438	49.438	14.831	49.438	49.438
2	8.295	27.651	77.089	8.295	27.651	77.089
3	2.563	8.543	85.632	2.563	8.543	85.632
4	1.479	4.929	90.561	1.479	4.929	90.561
5	1.188	3.961	94.522	1.188	3.961	94.522

Extraction Method: Principal Component Analysis.

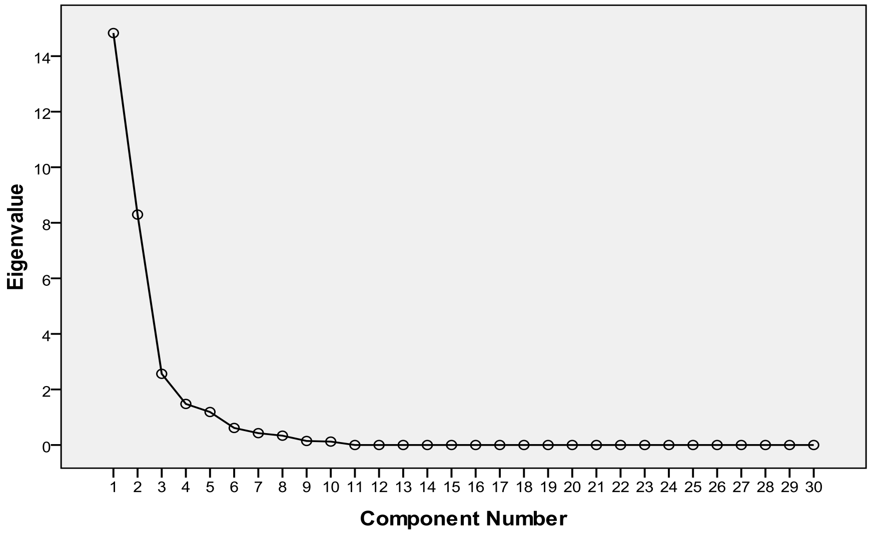


Figure 7-18 Scree plot for factor analysis

7.3.4 Task Outcomes

For each task, there were three possible outcomes: (1) the task was not attempted; (2) the participant successfully completed the task, and (3) the participant failed to complete the task by either thinking that they were successful when in actuality they failed, or by giving up while performing the task. Figure 7-19 shows the proportion of participants who attempted, succeeded or failed the task with each of the consumer products.

Of the 16 participants who performed the clock radio task, 56% successfully completed the task, and of the 16 participants who performed the mobile phone task, 19% completed it successfully. Of the 19 participants who performed the blender task, 100% successfully completed the task, and of the 18 participants who performed the vacuum cleaner task, 100%

completed it successfully. Thus the mobile phone task and the clock radio task had the highest and second highest failure rate respectively. This is summarised in Table 7-10.

Table 7-10 Proportion of participants performing, successfully completing and giving up during tasks

Product	Performed Task	Successfully Completed	Failed/Gave Up
Clock radio	16 of 19 participants (84%)	9 of 16 participants (56%)	7 of 16 participants (44%)
Mobile	16 of 19 participants (84%)	3 of 16 participants (19%)	13 of 16 participants (81%)
Blender	All 19 participants (100%)	19 participants (100%)	0 of 19 participants (0%)
Vacuum cleaner	18 of 19 participants (95%)	18 participants (95%)	0 of 18 participants (0%)

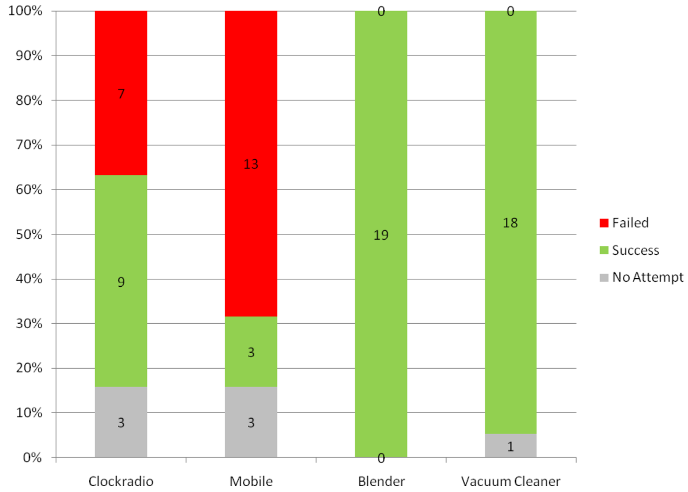


Figure 7-19 Graph of proportion of participants attempting, failing and being successful with each of the four tasks

Participants 8, 15 and 17 could not perform the clock radio and mobile tasks due to visual problems. Of the 7 participants who failed the clock radio task, 6 failed by giving up when they couldn't achieve the goal, while the remaining participant thought she achieved the goal when in actuality she did not. Of the 13 participants who failed the mobile phone task, all gave up after trying to reach the goal.

7.3.4.1 Mean Rated Difficulty for Each Task

Mean difficulty and frustration ratings were plotted for each product and compared. Figure 7-20 shows the mean ratings for the clock radio task. For this task, the most difficult actions were cognitive actions such as figuring out how to start the task and how to move on to the next action. Seeing the text 'PM' on the clock radio screen was also rated as being relatively difficult. Overall mental demand was rated higher than physical demand.

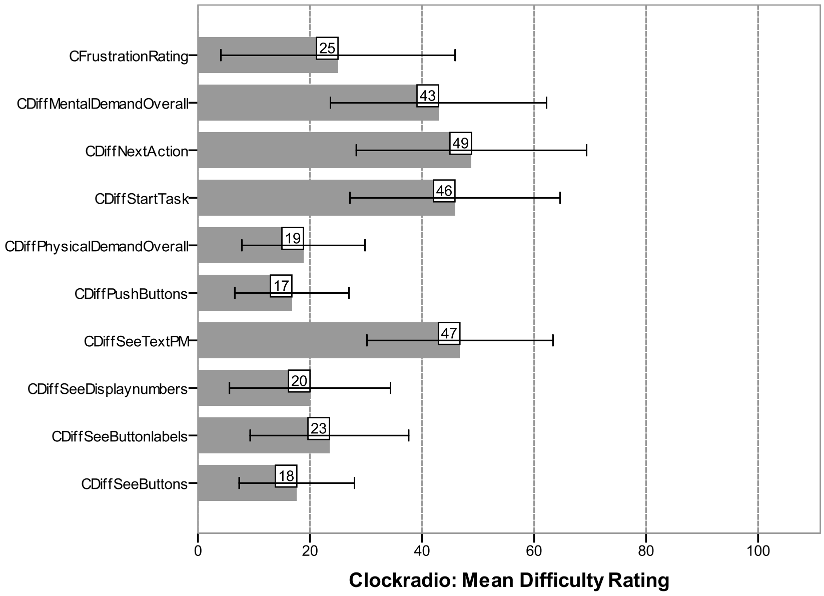


Figure 7-20 Mean difficulty ratings for clock radio task (error bars 95% confidence interval)

Figure 7-21 shows the mean ratings for the mobile phone task. As for the clock radio task, participants rated cognitive actions such as figuring out how to start the task and how to move on to the next action much higher than all other actions. Overall mental demand was also rated quite high in relation to overall physical demand. Seeing the text and the buttons on the phone also proved difficult for participants, and participants were fairly frustrated after performing this particular task.

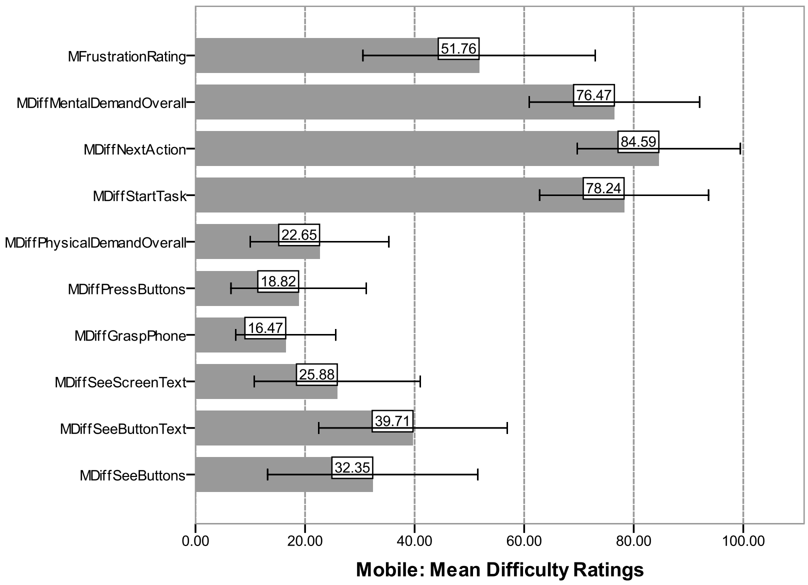


Figure 7-21 Mean difficulty ratings for mobile phone task (error bars 95% confidence interval)

Figure 7-22 shows the mean ratings for the blender task. In this task, the physical actions of opening/closing the cover of the blender were rated the most difficult. Lifting and pouring from the glass jug were also difficult motor actions. Overall physical demand ratings were higher than overall mental demand ratings, but not by a significant amount.

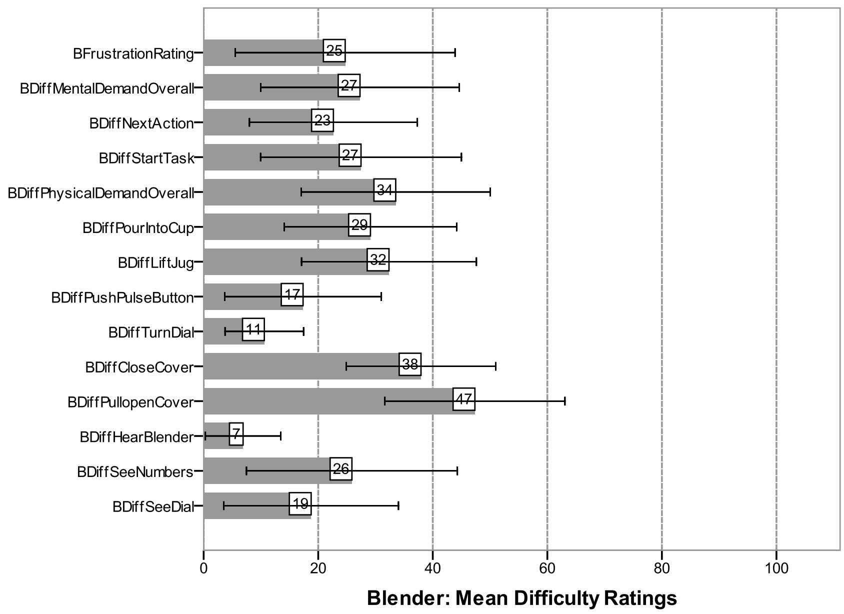


Figure 7-22 Mean difficulty ratings for blender task (error bars 95% confidence interval)

Figure 7-23 shows the mean ratings for the vacuum cleaner task. Figuring out how to start the task was rated the most difficult action, as participants found difficulty finding the power switch on the vacuum cleaner. Figuring out the next action in the sequence of operating actions and overall mental demand were also rated as being more difficult than the other actions. Participants rated overall physical demand, including pushing and pulling the vacuum cleaner, as the most difficult physical actions in the task. The vacuum cleaner also posed visual challenges to users in seeing the buttons and text on the chassis.

Across all products, the mobile phone had the highest mean ratings for difficulty in starting the task (M=78.24, SD=30.00), difficulty in working out subsequent actions (M=84.59, SD=28.96), and overall mental demand (M=76.47, SD=30.25). The mobile phone also had the highest mean rating for frustration experienced during the task (M=48.89, SD=41.82). In terms of visual demands, the small text on the clock radio display was rated the most difficult to see (M=52.37, SD=34.78), followed by seeing the numbers on buttons (M=46.05, SD=36.84) and seeing the actual buttons (M=39.47, SD=41.16) on the mobile phone.

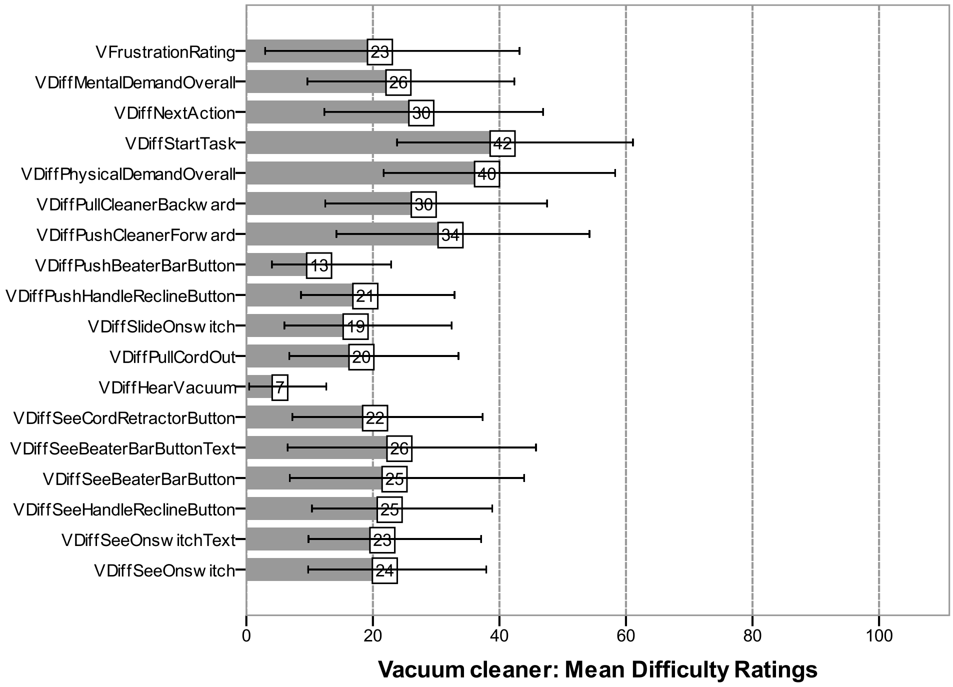


Figure 7-23 Mean difficulty ratings for vacuum cleaner task (error bars 95% confidence interval)

The physical actions of opening (M=47.37, SD=30.25) and closing (M=38.68, SD=26.03) the blender cover and pushing the vacuum cleaner forward (M=28.06, SD=30.69) were also rated as being the most difficult actions. In terms of overall mental demands, the mobile phone ranked the highest (M=76.47, SD=30.25), followed by the clock radio (M=42.94, SD=37.54), the blender (M=28.11, SD=32.58) and the vacuum cleaner (M=27.94, SD=27.60). For mean frustration ratings, the mobile phone once again ranked the highest (M=48.89, SD=41.82), followed by the vacuum cleaner (M=28.33, SD=36.22), the clock radio (M=26.39, SD=39.91) and the blender (M=22.11, SD=36.03). The mobile phone provided the greatest cognitive challenge, while the vacuum cleaner provided the greatest physical challenge to participants. The error bars on all the four graphs of mean difficulty ratings were quite large, possibly due to the small sample size of the study. The results therefore can be interpreted as indicative, with a larger sample size being required for statistically significant results.

7.3.4.2 Times and Errors

Histograms of task times (in seconds) are shown in Figure 7-24 for each product. The vacuum cleaner has the longest average time (M=169.94, SD=77.11), followed by the blender (M=140.89, SD=86.12), then the mobile phone (M=128.31, SD=99.59) and finally the clock radio (M=93.56, SD=58.26).

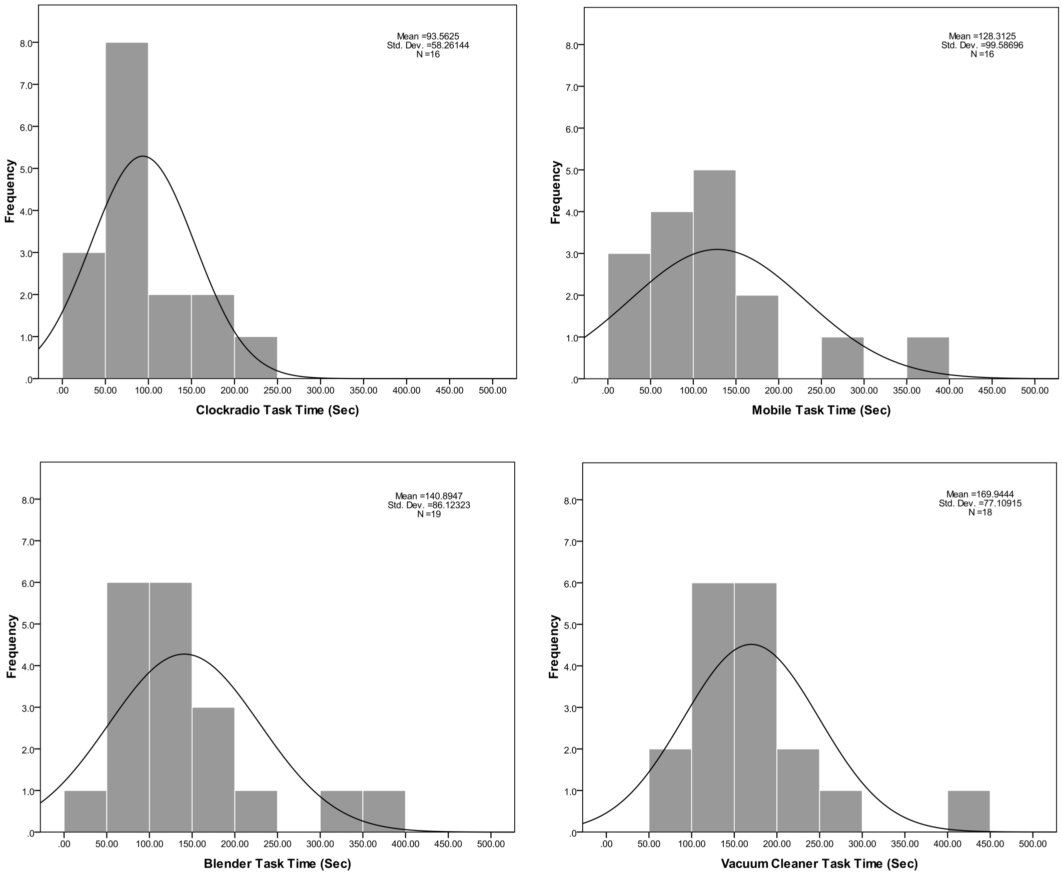


Figure 7-24 Histograms of task times for each of four products

The clock radio and mobile phone tasks had the highest failure rate, and participants in general gave up quickly after finding the task too difficult. This explains the shorter mean task times. Figure 7-25 shows histograms of task errors for each product. The mobile phone recorded the highest mean errors (M=3, SD=2.63), followed by the clock radio (M=1.5, SD=0.89), the blender (M=1.26, SD=1.41), and finally the vacuum cleaner (M=1.11, SD=1.32). The high frequency of one error for the clock radio task was due to participants not being able to figure out that a ‘time set’ button must be held in order to set the time, and thus entering a trial and error exploration of the product interface. This acted as a barrier for successfully completing the task. Similarly, the high frequency of one error for the mobile phone task is due to some participants not being able to figure out the correct menu option once the phone menu had been entered. The random exploration that ensued culminated in task failure. Designers therefore need to be aware of introducing these barriers in product design where ‘hold and set’ strategies could set up an insurmountable barrier to task completion.

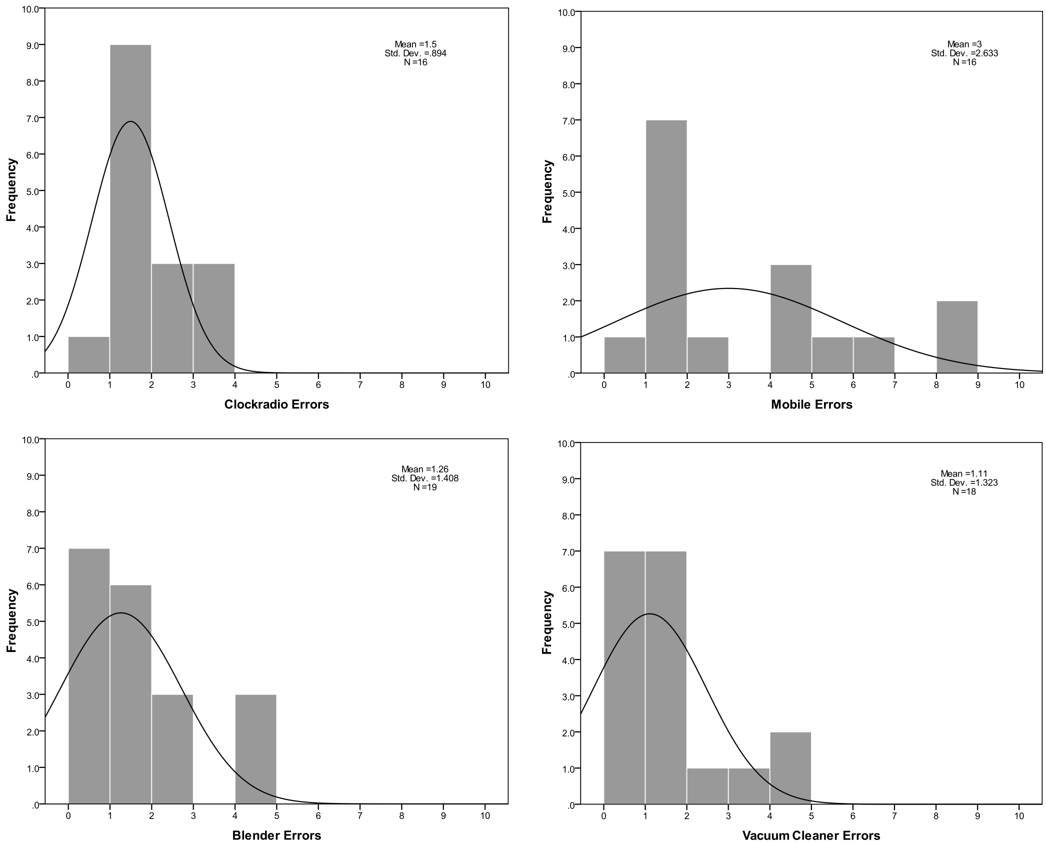


Figure 7-25 Histograms of task errors for each of four products

Figure 7-26 shows scatter plots of task time versus task errors for each of the four products used in the study. For the mobile, blender and vacuum cleaner, the graphs demonstrate a positive relationship i.e. the task time increases as the number of errors increase. The mobile phone graph shows the best linear relationship of the three with $r(14)=0.726$, $p< 0.01$. It indicates that in these cases, there are no ‘speed versus accuracy’ tradeoffs occurring where faster task times lead to more errors and vice versa. The clock radio scatter plot shows a weak, statistically insignificant negative relationship: $r(14)=-.149$. Participants making one error in this task demonstrated a large variation in task time which is responsible for skewing the graph. Given the distribution of the other points, it can be assumed that there is no ‘speed versus accuracy’ tradeoff occurring in the clock radio task, despite the negative correlation coefficient.

Participants 11 (retired technology consultant) and 13(retired army officer) tended to perform well in the product tasks with relatively low task times and errors. Participant 11 was also the only participant to succeed on all product tasks. This may have been possible due to their previous experience and training with technology.

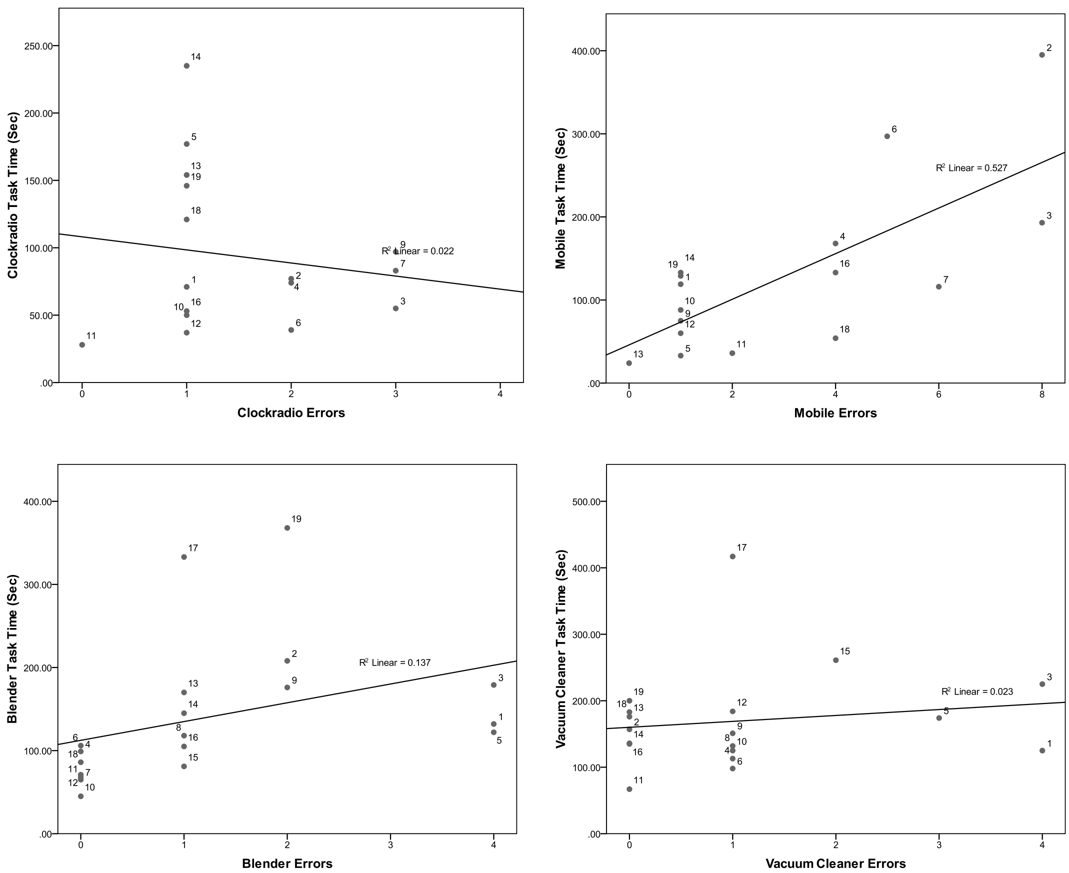


Figure 7-26 Scatter plots of task time versus errors made for the four product tasks

7.3.5 Capability-Demand Analysis

In the following sections, an analysis of the relationship between measured user capability and measured product demand will be presented for sensory, cognitive and motor actions for all four products. This analysis will take the form of capability-demand graphs as illustrated in Figure 7-27. In these graphs, measured user capability will be plotted on the horizontal axis for a given action. Rated difficulty for the action will be plotted on the vertical axis. The particular demand on user capability will be plotted as a red ‘demand’ line on the graph as shown on the bottom graph of the figure.

This presumed inverse relationship between difficulty and capability is shown simplified on the top graph of Figure 7-27. It is expected that participants with low capability would have high difficulty scores, while participants with high capability should have low difficulty scores (in relation to a demand line). If points appear in the quadrant described as high capability-high difficulty, it could mean that the difficulty scores are being inflated due to factors such as varied use of the rating scale, frustration buildup during the task, multidimensional judgments and even hidden factors such as pain. If points appear in the quadrant described as low capability-low difficulty, it may be due to factors such as the use of coping strategies (substituting one capability for another) or a threshold of demand below which the action would be relatively easy despite low capability levels.

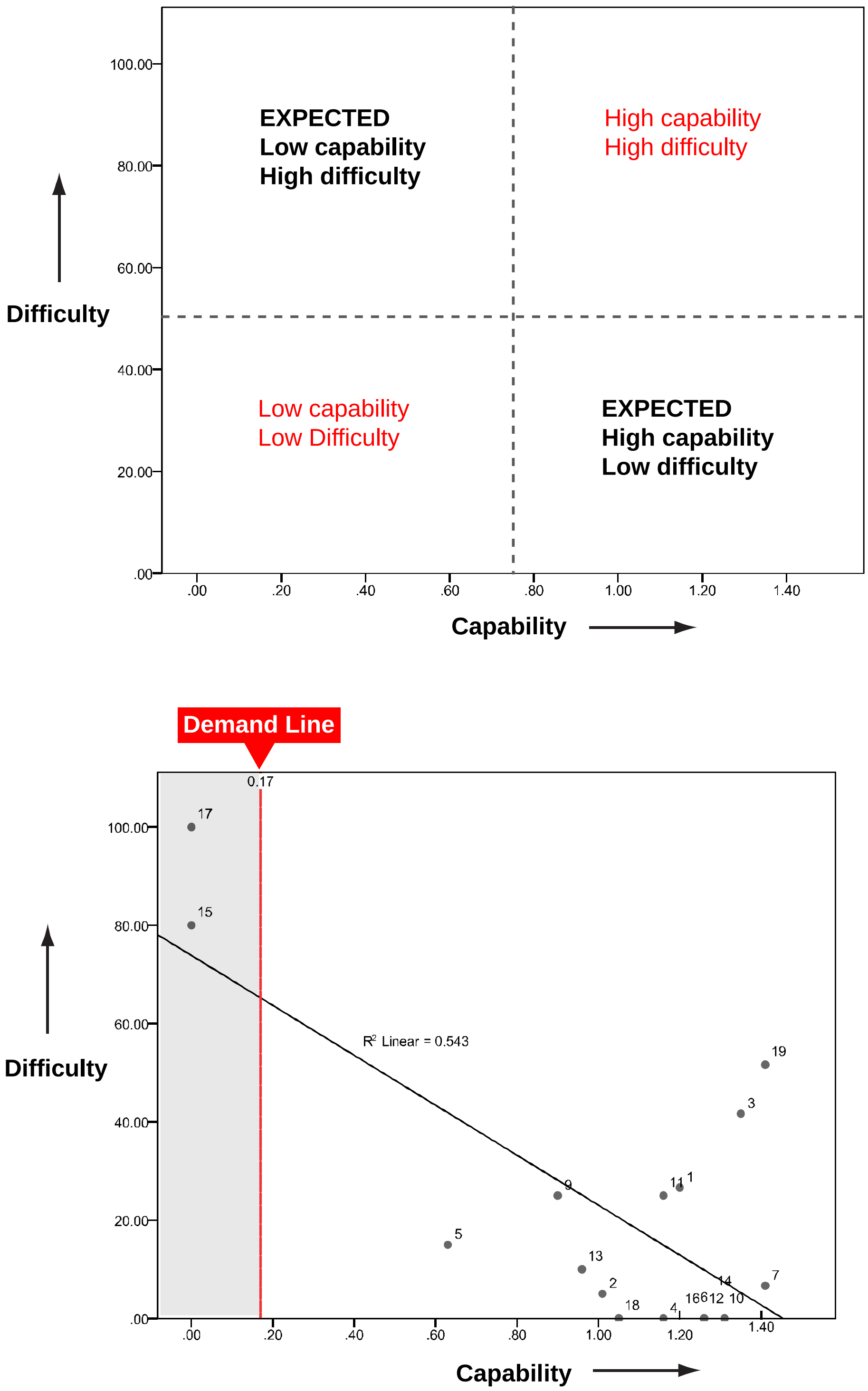


Figure 7-27 Capability-Demand generic graph template for interpreting results

For example, as seen in the bottom graph of Figure 7-27, if the capability in question is pulling strength, the demand line would be the exact force that the user would be required to pull on the product. It is expected that as one moves to the right of the red demand line, experienced difficulty scores should fall. As one moves to the left of the red demand line, rated difficulty should be near the maximum, i.e. it should be exceedingly difficult or near impossible to perform the action. In other words, as capability increases, difficulty ratings are expected to decrease in an inverse relationship.

Product demand values were calculated from measurements taken from the four products. Table 7-11 shows calculated demand values for textual features and interface objects. In order to calculate these values, the average distance of use from the product was estimated. The visual angle subtended on the retina was calculated (LogMar demand) using the height and viewing distance of the text and objects and the formula $\tan^{-1}(H/2D)$ converted to degrees (Colenbrander, 2003; Fletcher & American Academy of Ophthalmology., 1999). The Log contrast sensitivity demand for each feature was calculated from the estimated contrast values measured with the contrast cards.

Table 7-11 Product demand values for visual features

Visual Demands	Product	Feature	Text Height, H, (mm)	Diameter, D, (mm)	Contrast (%)	Distance (mm)	LogMar Demand	Log Contrast Sensitivity Demand
button text labels	Clock radio	Text	1.5	-	50	400	0.411	0.30
numbers on the digital display	Clock radio	Text	14	-	14	400	1.38	0.85
text ‘PM’ on the display	Clock radio	Text	2	-	33	400	0.536	0.48
the text and numbers on the buttons	Mobile	Text	2	-	67	400	0.536	0.17
text on the display screen	Mobile	Text	2	-	33	400	0.536	0.48
numbers around the control dial	Blender	Text	3	-	67	600	0.536	0.17
text on the *on/off* switch	Vacuum	Text	3	-	60	600	0.536	0.22
text label for the beater bar button	Vacuum	Text	3	-	67	1000	0.314	0.17
buttons that you used	Clock radio	Object		1.7	50	400	1.164	0.30
buttons	Mobile	Object		5	60	400	0.633	0.22
blender control dial	Blender	Object		48	67	600	1.439	0.17
cord pull/retractor button	Vacuum	Object		19	25	600	1.037	0.60
on/off switch	Vacuum	Object		14	43	600	0.904	0.37
handle recline button	Vacuum	Object		25	25	1000	1.156	0.60

Table 7-12 shows measured demand values (forces) for activating or manually handling interface features on the four products. These included weights, linear forces and rotational forces. The product demand values were plotted as red vertical lines on capability/demand graphs as elaborated in the next sections.

Table 7-12 Product demand values for motor actions

Motor Demands	Product	Action Type	Force/Torque
Push Buttons	Clock Radio	Finger	3.3 N
Push buttons	Mobile	Finger	2.9 N push force
turn the control dial	Blender	Finger	0.236 Nm clockwise, 0.177Nm anticlockwise
push the *PULSE* button	Blender	Finger	9N
slide the on/off switch	Vacuum Cleaner	Finger	13 N
Lift chassis	Mobile	Grasping	0.82 N weight
Open cover of the blender jug	Blender	Grasping	19.8 N
lift the blender jug	Blender	Grasping	19.3 N
pour from the blender jug into the cup	Blender	Grasping	19.3 N
pull the plug and cord out	Vacuum Cleaner	Grasping	14.7 N
push the cleaner forward	Vacuum Cleaner	Grasping	15 N
pull the cleaner backward	Vacuum Cleaner	Grasping	30 N

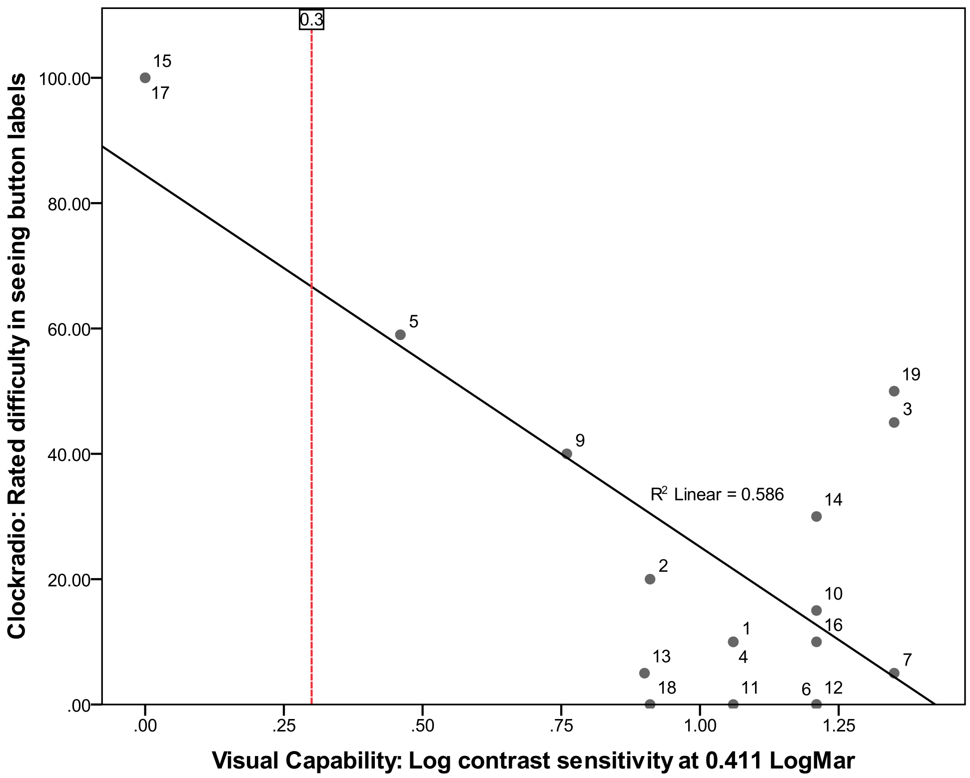
7.3.6 Sensory Capabilities and Demands

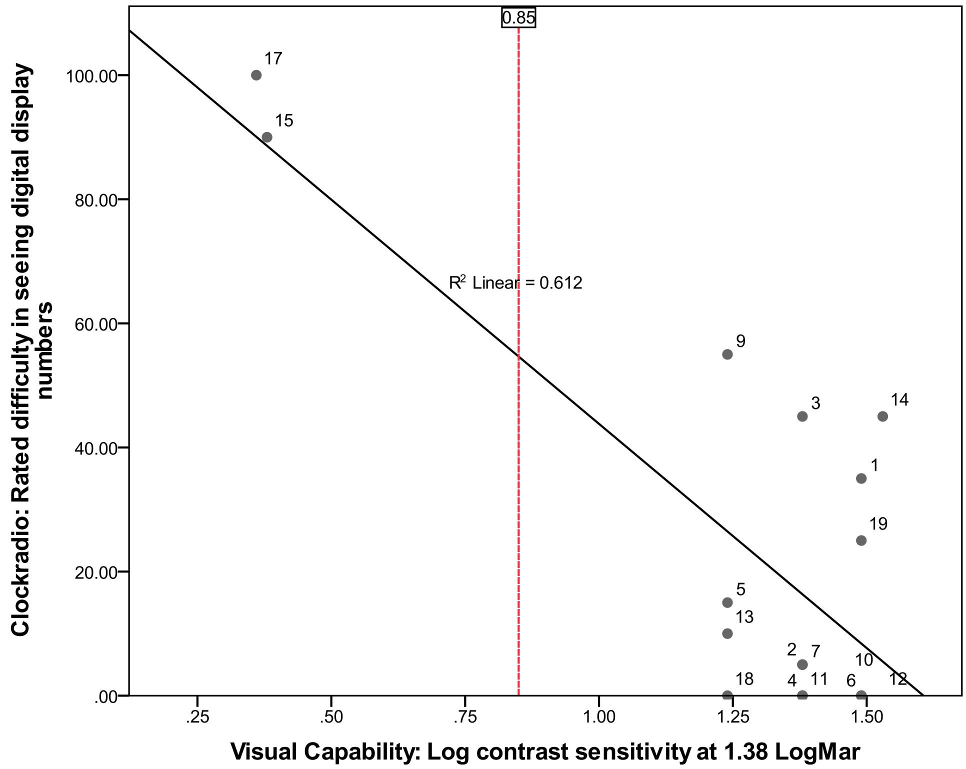
Scatter plots between visual capabilities and rated difficulty in visual actions were generated for seeing textual features on the products and seeing product controls.

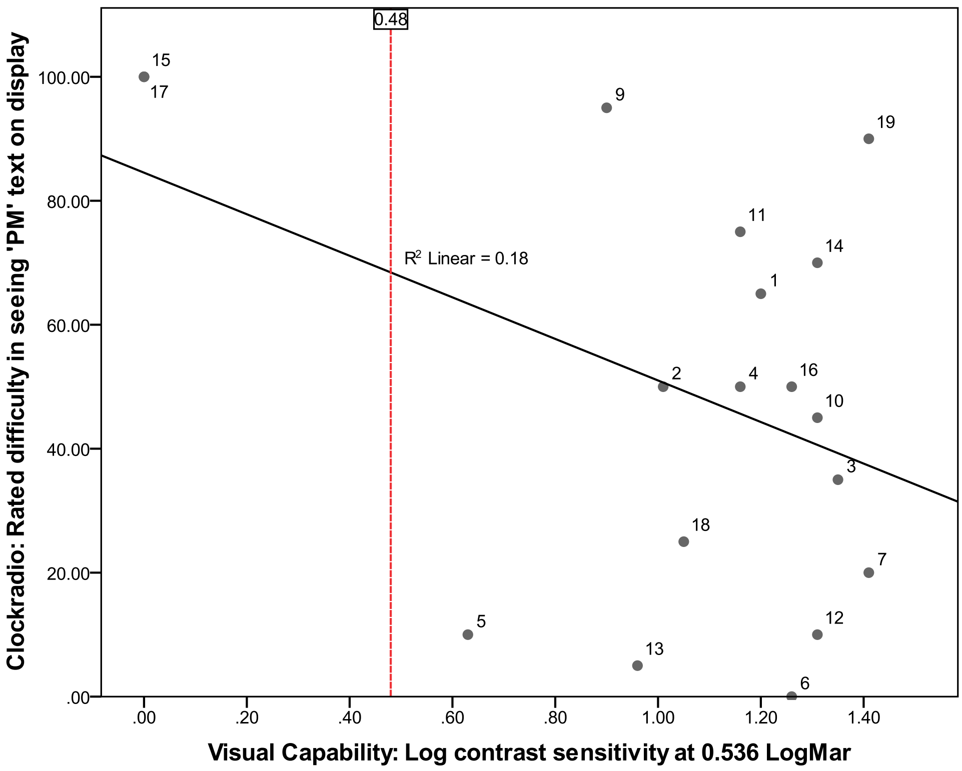
7.3.6.1 Seeing Textual Features

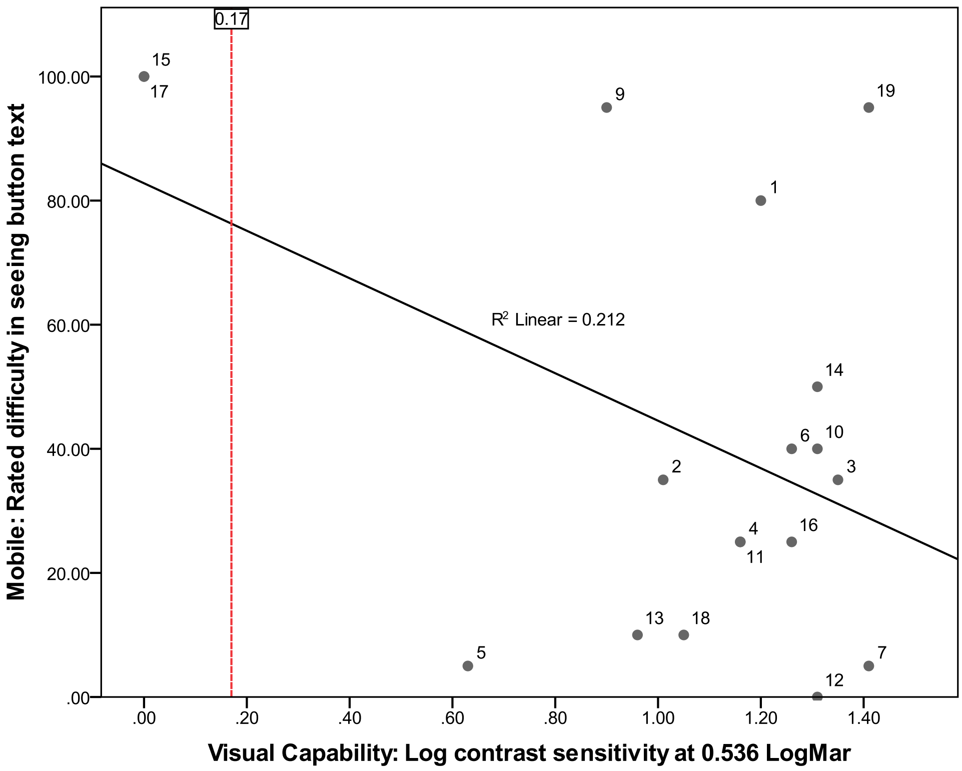
Figure 7-28 over the next four pages shows 8 scatter plots relating textual features to visual capability.

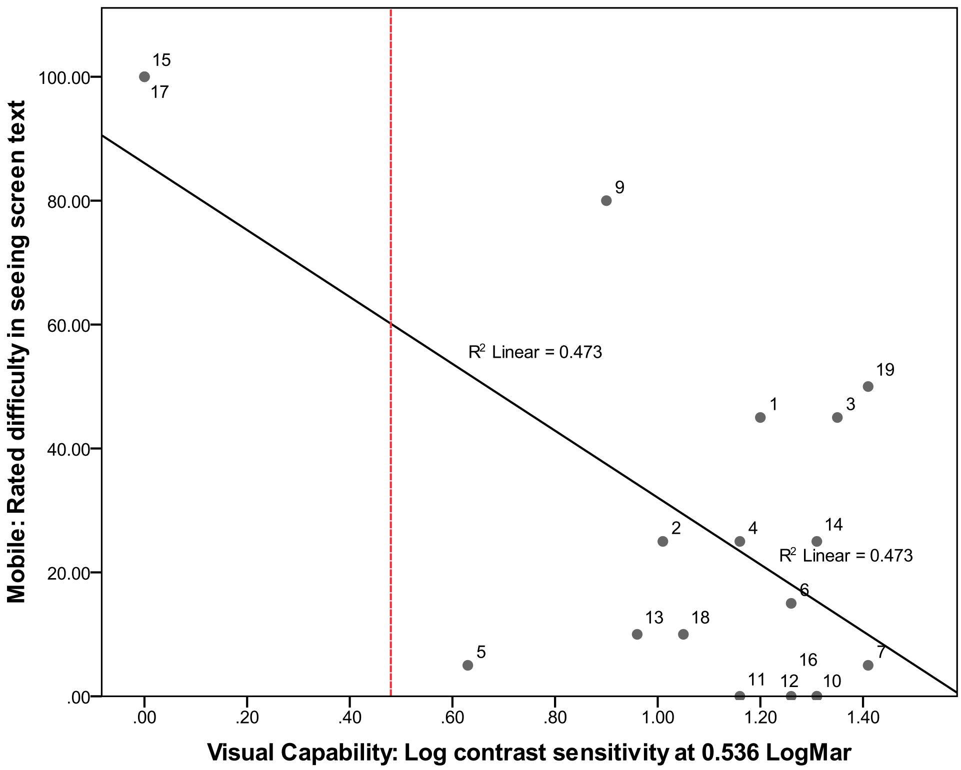
Table 7-13 gives the values of the correlation coefficients (Pearson's r) for each action. Values in this table (and all following tables with correlation coefficient values) are highlighted in red if the absolute value of r is strong (0.7 to 1) and in yellow if moderate (0.4 to 0.7). Values that are not highlighted are to be interpreted as weak (0 to 0.4). Each graph plots the contrast capabilities of participants at a particular size measured in LogMar against the rated difficulty in seeing text of a given contrast (at the same size). For example, the contrast demand of the clock radio digital display text is 0.85 (in Log Contrast Sensitivity units) at the size of 1.38 LogMar. A linear model accounted for a significant amount of the variance for actions involving reading textual features on the products.

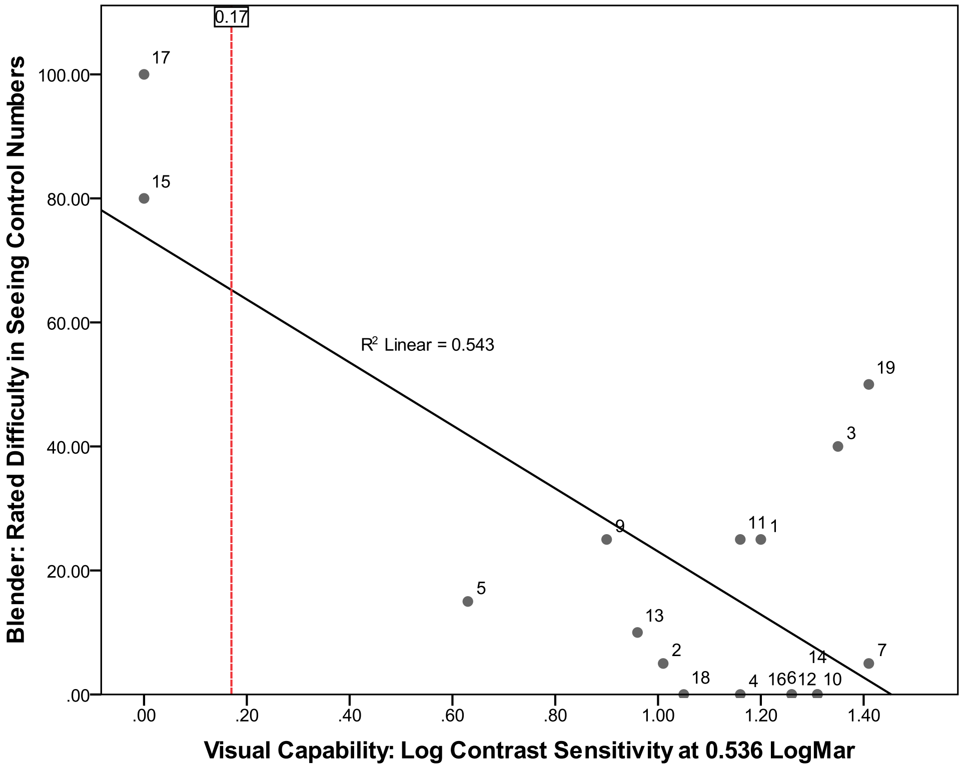

0.3
Clockradio: Rated difficulty in seeing button labels
Visual Capability: Log contrast sensitivity at 0.411 LogMar
R² Linear = 0.586
100.00
80.00
60.00
40.00
20.00
.00
.00
.25
.50
.75
1.00
1.25

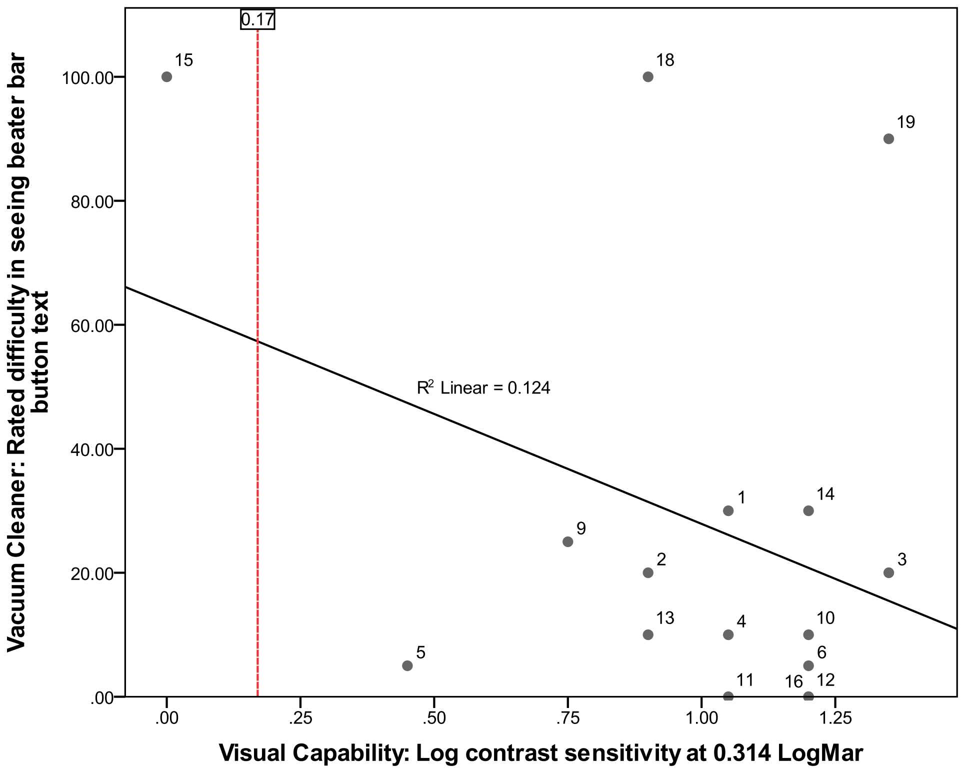

0.85
Clockradio: Rated difficulty in seeing digital display numbers
Visual Capability: Log contrast sensitivity at 1.38 LogMar
R² Linear = 0.612
100.00
80.00
60.00
40.00
20.00
.00
.25
.50
.75
1.00
1.25
1.50


0.48
R² Linear = 0.18
Clockradio: Rated difficulty in seeing 'PM' text on display
Visual Capability: Log contrast sensitivity at 0.536 LogMar
.00
20.00
40.00
60.00
80.00
100.00
.00
.20
.40
.60
.80
1.00
1.20
1.40


0.17
R² Linear = 0.212
Mobile: Rated difficulty in seeing button text
Visual Capability: Log contrast sensitivity at 0.536 LogMar
.00
20.00
40.00
60.00
80.00
100.00
.00
.20
.40
.60
.80
1.00
1.20
1.40


Mobile: Rated difficulty in seeing screen text
Visual Capability: Log contrast sensitivity at 0.536 LogMar
R² Linear = 0.473
R² Linear = 0.473


0.17
Blender: Rated Difficulty in Seeing Control Numbers
Visual Capability: Log Contrast Sensitivity at 0.536 LogMar
R² Linear = 0.543



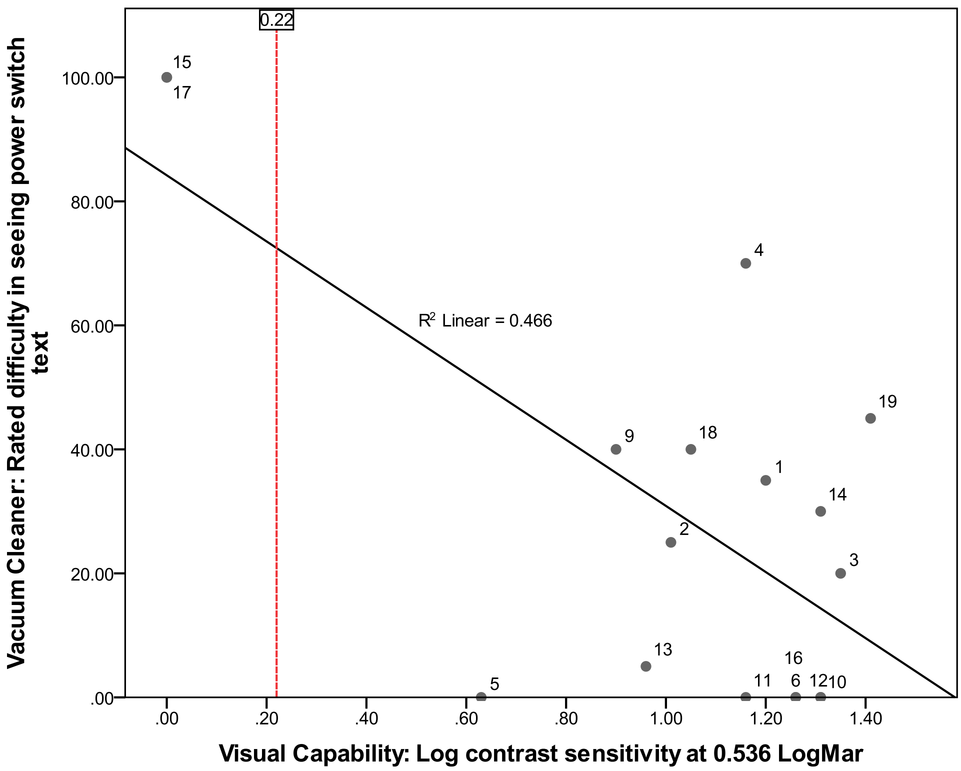


Figure 7-28 Eight capability-demand graphs for seeing textual features on the four products

Table 7-13 Correlation coefficients for seeing textual features

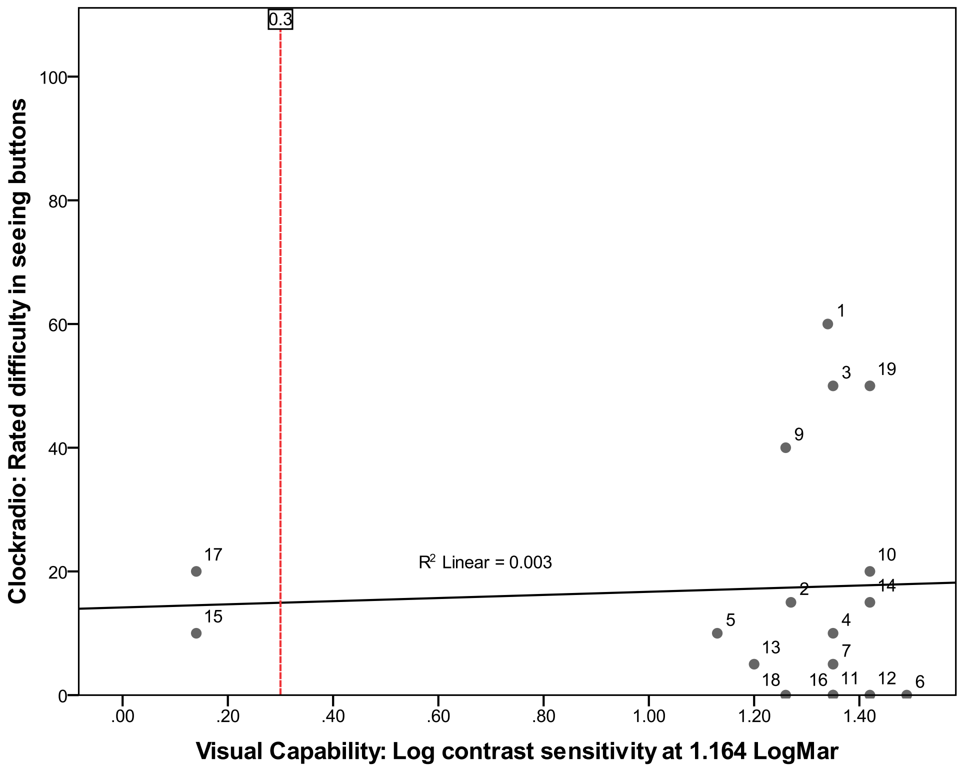
Product	Action	Correlation Coefficient
Clock	Seeing numbers on the digital display	-0.782
Clock	Seeing button labels	-0.766
Blender	Seeing numbers on the rotary control	-0.737
Mobile	Seeing texing on the screen	-0.688
Vacuum Cleaner	Seeing text on power switch	-0.683
Mobile	Seeing button text	-0.461
Clock	Seeing 'PM' text on the display	-0.424
Vacuum Cleaner	Seeing beater bar button text	-0.352

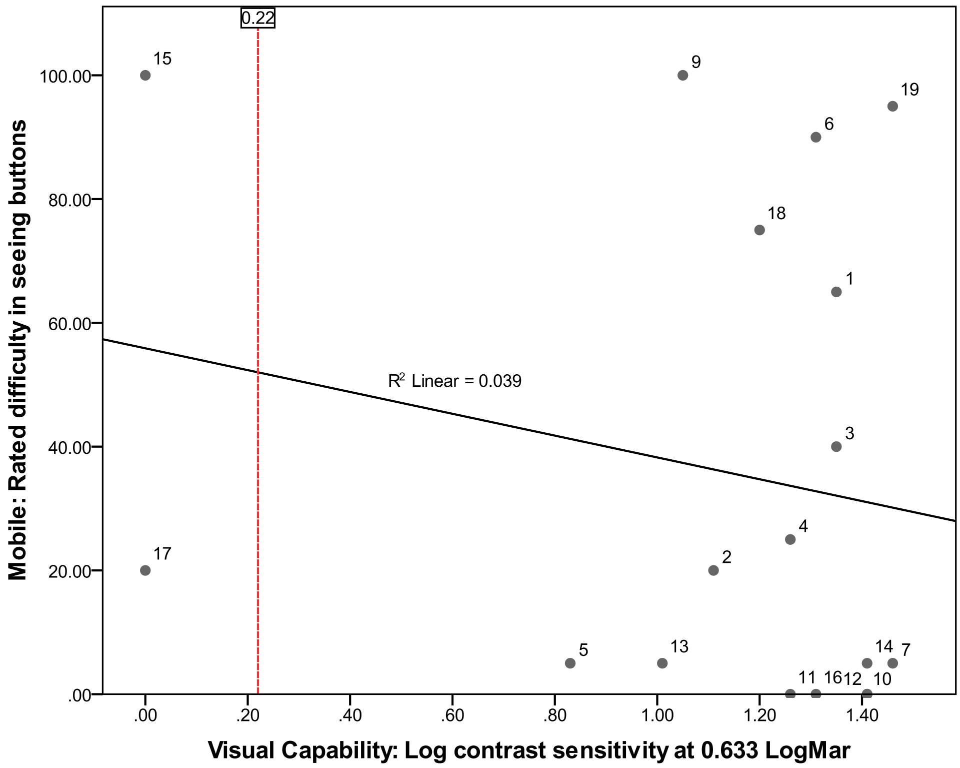
Figure 7-28 shows fairly strong negative relationships for seeing button labels r(16)=-0.766, p< 0.01 and reading numbers on the digital display r(16)=-0.782, p< 0.01 of the clock radio. Strong relationships were also observed for seeing the screen text on the mobile r(16)=-0.688, p< 0.01, seeing the control numbers on the blender r(16)=-0.737, p< 0.01, and seeing the power switch text on the vacuum cleaner r(15)=-0.683, p< 0.01. However, other relationships showed weaker correlations including seeing the 'PM" text on the clock radio display r(16)=-0.424, p< 0.05, seeing the button text on the mobile r(15)=-0.461, p< 0.05 and seeing the beater bar button text on the vacuum cleaner r(14)=-0.352.

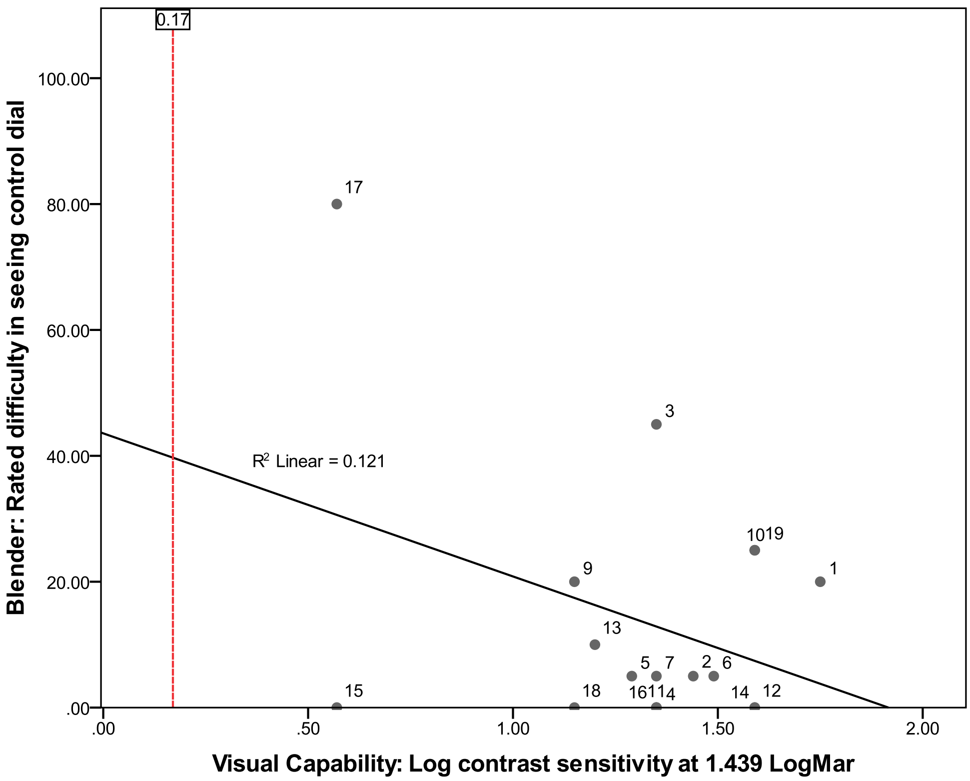
In general, participants 1, 3, 9 and 19 tended to score difficulty much higher than other participants. Participants 1, 3 and 9 had comparatively low cognitive scores compared to the rest of the sample, which may be a factor in the way they used the difficulty scale to make judgements. Participant 19's cognitive scores, however, were not as low as 1, 3 and 9. The graphs show, in general, a vertical spread to difficulty ratings at high levels of capability. This indicated that participants had different levels of difficulty even though their capability levels were roughly similar. However, the trend of correlation lines indicate that rated difficulty does increase as a participant's capability threshold approaches the product demand line. In addition, participants whose capability fell to the left of the demand line rated difficulty as being maximum or near maximum.

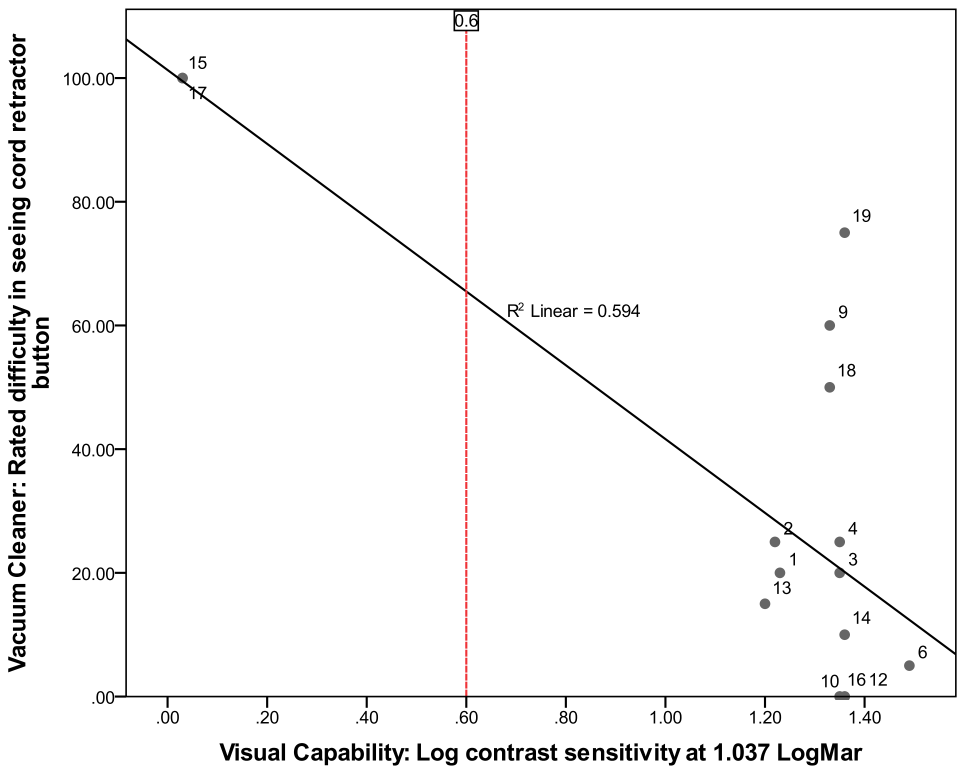
7.3.6.2 Seeing Product Controls and Features

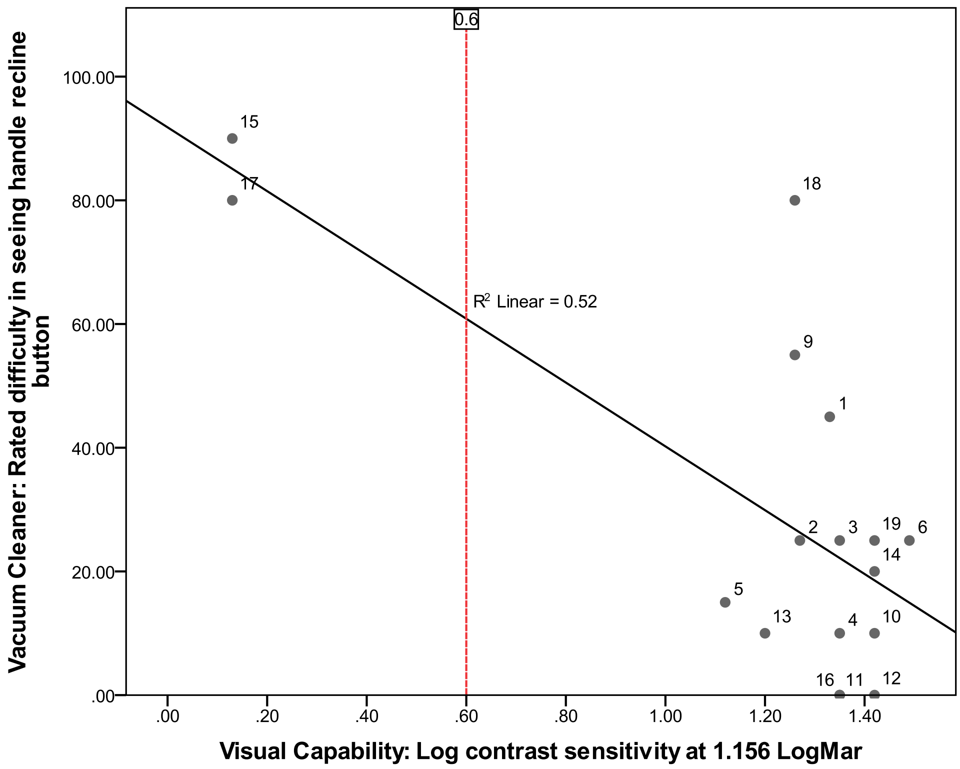
Some significant linear relationships were also found for actions involving seeing product controls and features, as shown in the 6 graphs of Figure 7-29 over the next three pages and in Table 7-14. Seeing the cord retractor button r(14)=-0.771, p< 0.01, the handle recline button r(15)=-0.721, p< 0.01 and the power switch r(15)=-0.749, p< 0.01 on the vacuum cleaner all produced fairly strong linear relationships.


0.3
R² Linear = 0.003
Clockradio: Rated difficulty in seeing buttons
Visual Capability: Log contrast sensitivity at 1.164 LogMar


0.22
R² Linear = 0.039
Mobile: Rated difficulty in seeing buttons
Visual Capability: Log contrast sensitivity at 0.633 LogMar


0.17
R² Linear = 0.121
Blender: Rated difficulty in seeing control dial
Visual Capability: Log contrast sensitivity at 1.439 LogMar
100.00
80.00
60.00
40.00
20.00
.00
.00
.50
1.00
1.50
2.00


0.6
R² Linear = 0.594
Vacuum Cleaner: Rated difficulty in seeing cord retractor button
Visual Capability: Log contrast sensitivity at 1.037 LogMar
100.00
80.00
60.00
40.00
20.00
.00
.00
.20
.40
.60
.80
1.00
1.20
1.40



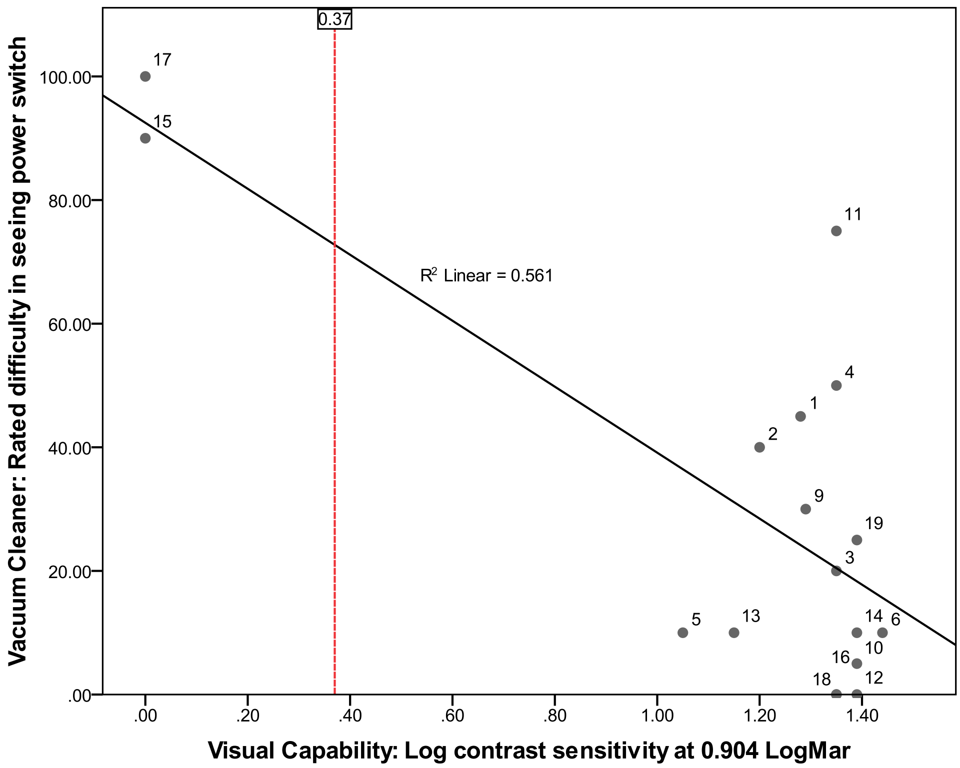


Figure 7-29 Six capability-demand graphs for seeing interface features on the four products

However, other cases showed no significant linear relationships, for example seeing the buttons on the clock radio r(16)=0.051, seeing the buttons on the mobile r(16)=-0.197 and seeing the control dial on the blender r(16)=-0.197. Similar to the graphs in the last section for seeing text, participants 1, 3, 9 and 19 tended to score difficulty much higher than other participants. The graphs also show a similar vertical spread to difficulty ratings at high levels of capability. Unlike the last three graphs in the figure for the vacuum cleaner, the first two graphs show participants with low difficulty ratings, even though the points lie to the left of the demand line. The slope of the first graph is also positive with r(16)=0.051 given the vertical spread of difficulty scores. This may be due to the use of the contrast sensitivity measure for text as an approximation in place of a spatial contrast sensitivity test using sinusoidal gratings (Colenbrander, 2003; Fletcher & American Academy of Ophthalmology., 1999; Schiffman, 2000). It may be that the approximation holds for the vacuum cleaner when the distance of features from the user when in operation is approximately 1 – 1.5m. The perceived size of the feature would then be comparable with the stroke width of the text used in the capability measure.

Table 7-14 Correlation coefficients for seeing product controls and features

Product	Action	Correlation Coefficient
Vacuum Cleaner	Seeing cord retractor button	-0.771
Vacuum Cleaner	Seeing power switch	-0.749
Vacuum Cleaner	Seeing handle recline button	-0.721
Blender	Seeing control dial	-0.348
Mobile	Seeing buttons	-0.197
Clock	Seeing buttons	0.051

7.3.6.3 Within Product Comparisons

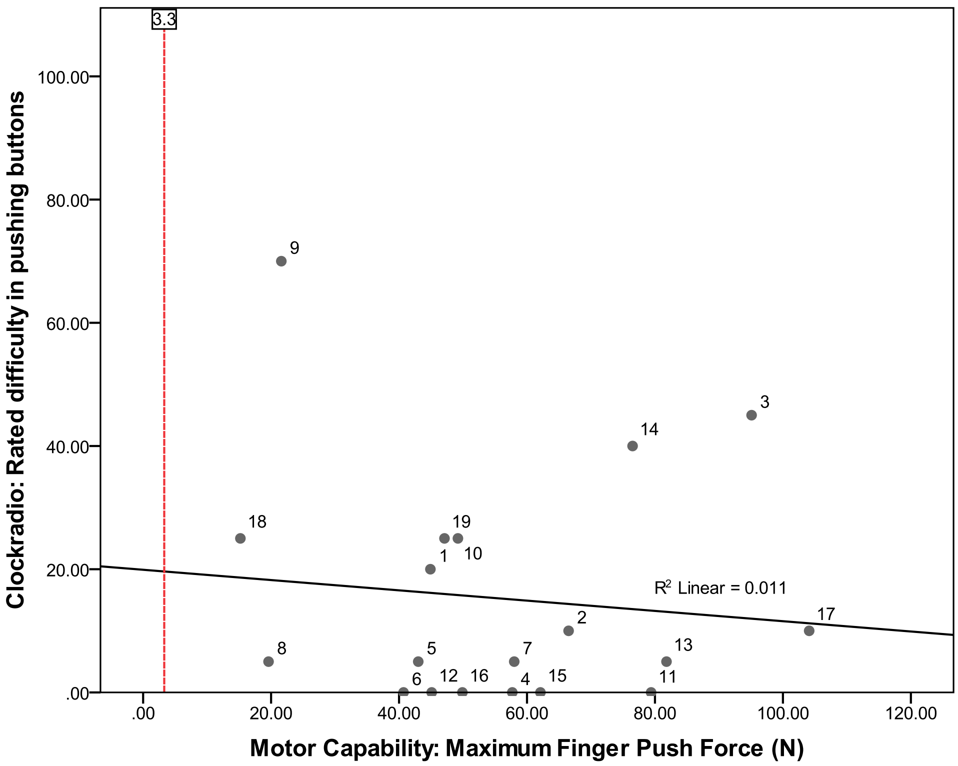
For the clock radio, the difficulty ratings for reading text on the LCD followed a similar pattern. There was a wide vertical spread of points indicating that different participants of similar capability levels were finding different degrees of difficulty with the task. The difficulty ratings for reading text on the clock radio buttons were less vertically spread out however, resulting in a better linear fit. The mobile showed the opposite pattern where text on the screen had less of a vertical spread of points than reading text on the button chassis. There was a difference between the clock radio LCD and the mobile phone display in that the clock radio LCD was light text on a dark background, while the mobile phone display was dark text on a light background. This may have been a factor in the perceived difficulty of reading on both screens. The blender and the vacuum cleaner followed the vertical distribution of points for visual actions within the task.

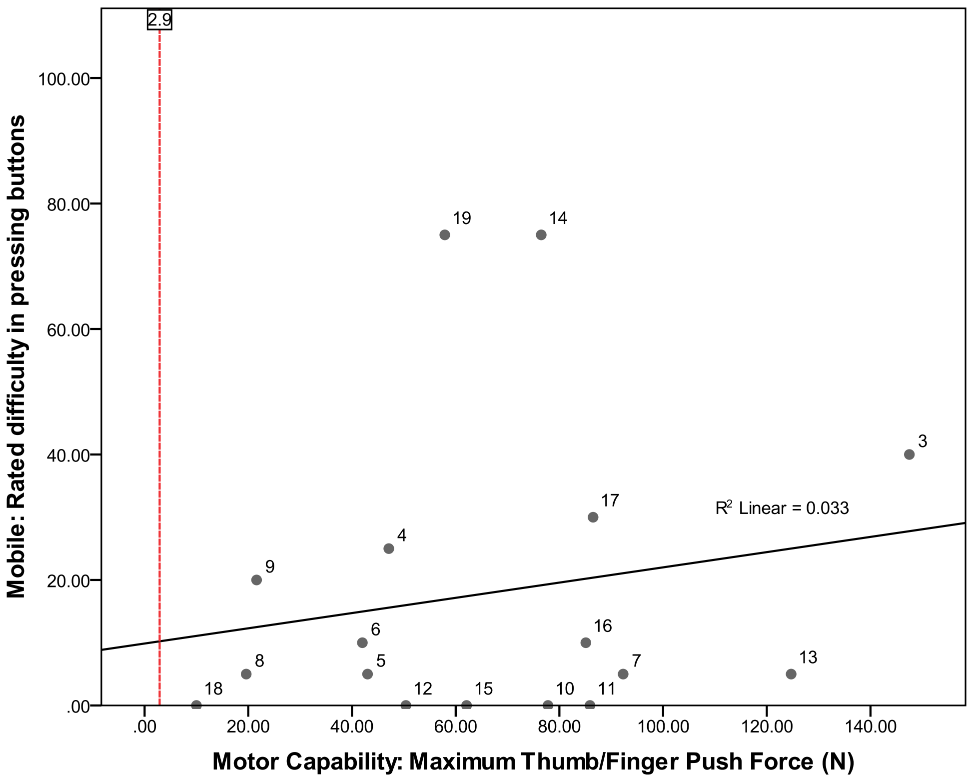
7.3.7 Motor Capabilities and Demands

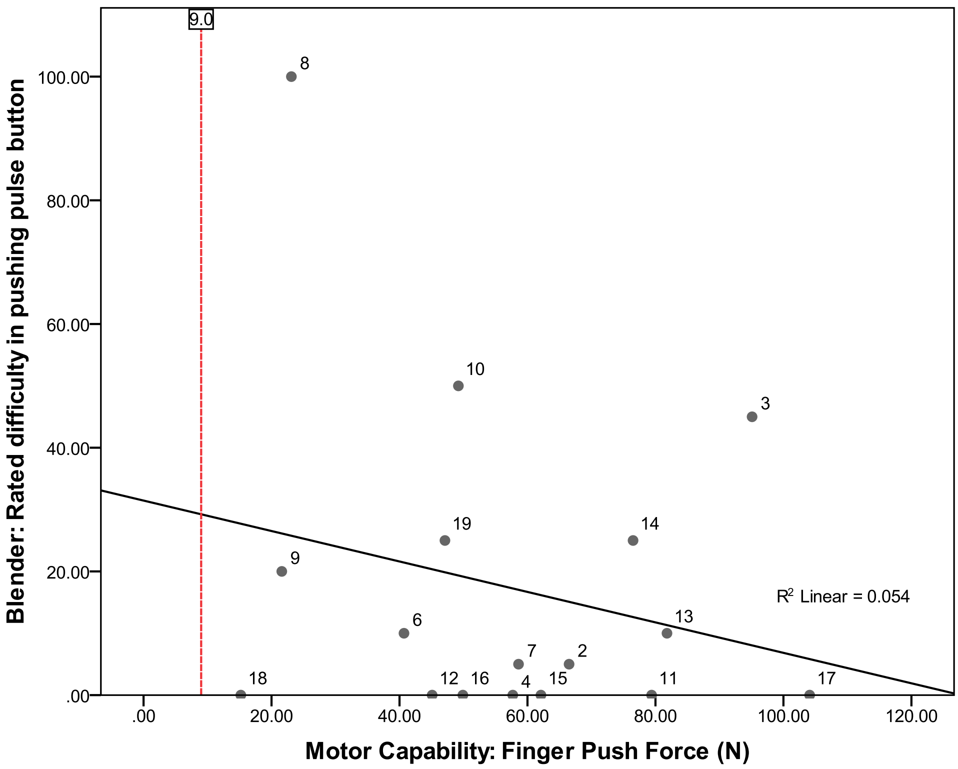
In this section, relationships between motor capabilities and product demands are presented. Physical actions for the 4 products were divided into actions involving finger forces and fine manipulations, grasping and large push/pull forces, and actions involving foot push forces. Though measures of both maximum and comfortable levels of force exertion were collected, only maximum values will be used in the following sections as some participants (especially those with low cognitive scores) had difficulty in understanding the difference between both measures. Therefore in some cases participants only performed the test for maximum force exertion. For each participant, the force capability of the actual hand used to perform the action was used (extracted from the participant's video). If both hands were used to perform a physical action, the maximum force capability of the stronger hand was used.

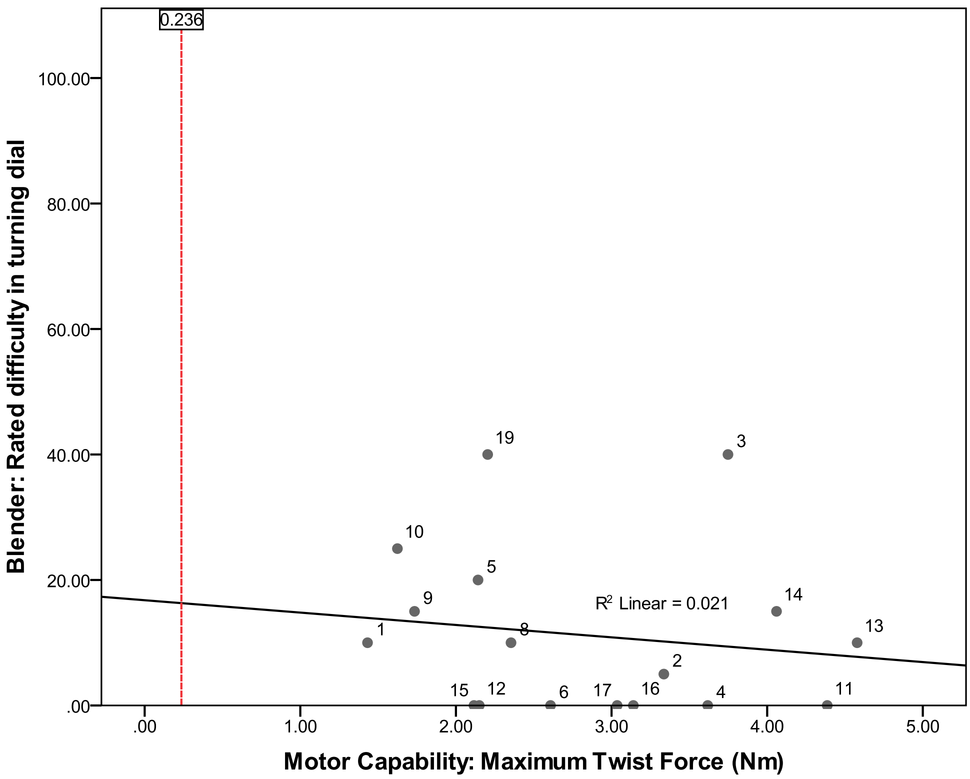
7.3.7.1 Actions Involving Small Finger Forces and Fine Manipulations

Figure 7-30 shows 5 graphs that involve actions utilising small finger forces and fine manipulations such as pressing buttons, rotating a control dial and sliding a switch. Correlation coefficients are given in Table 7-15. Pushing the clock radio buttons $r(17)=-0.105$, pressing buttons on the mobile phone $r(15)=0.181$, pushing the pulse button on the blender $r(15)=-0.232$, turning the dial on the blender $r(15)=-0.145$ and sliding the power switch on the vacuum cleaner $r(16)=-0.017$ all resulted in insignificant, weak linear relationships.


3.3
Clockradio: Rated difficulty in pushing buttons
Motor Capability: Maximum Finger Push Force (N)
R² Linear = 0.011


2.9
Mobile: Rated difficulty in pressing buttons
Motor Capability: Maximum Thumb/Finger Push Force (N)
R² Linear = 0.033


9.0
Blender: Rated difficulty in pushing pulse button
Motor Capability: Finger Push Force (N)
R² Linear = 0.054
.00
20.00
40.00
60.00
80.00
100.00
120.00


0.236
Blender: Rated difficulty in turning dial
Motor Capability: Maximum Twist Force (Nm)
R² Linear = 0.021
.00
1.00
2.00
3.00
4.00
5.00

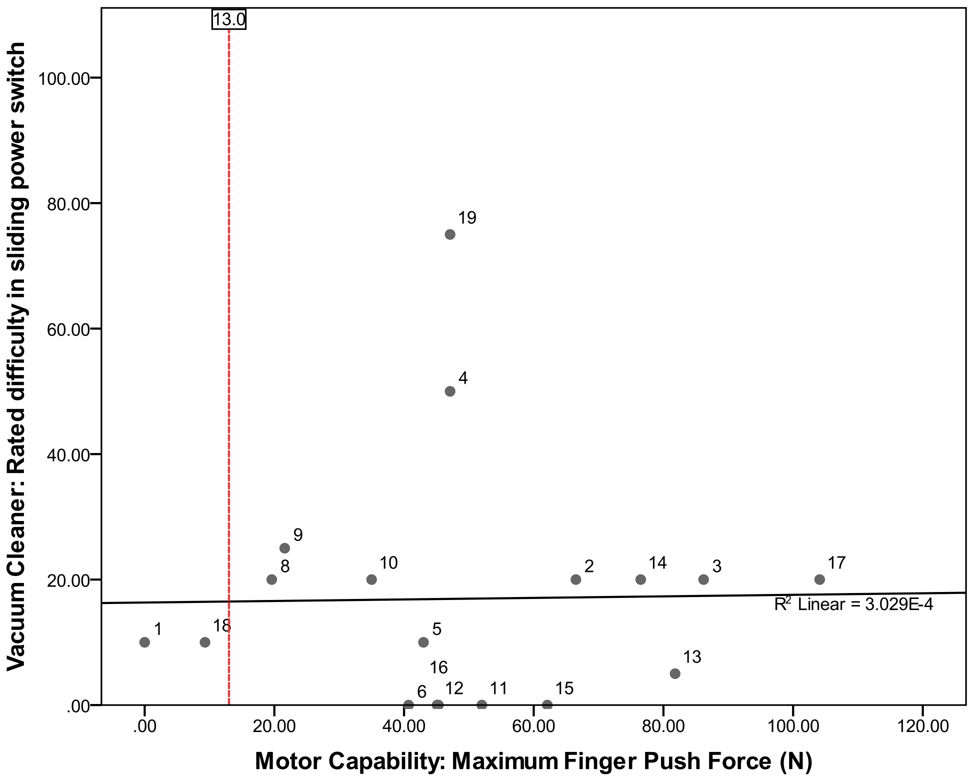


Figure 7-30 Five capability-demand graphs for finger forces and fine manipulations

Though a few participants rated increased difficulty as their maximum capability approached the demand line, a larger number of participants showed relatively low difficulty scores even as their maximum capability decreased. One possibility is that the demand values were particularly low for these actions, and at the scale of these small forces, users find the action relatively easy to perform, even though their measured capability is low. Even so, participants 3, 8, 9, 10, 14, and 19 tended to rate difficulty higher than the other participants in these actions. These participants all had conditions which affected their use of their hands such as arthritis, repetitive strain injury and deformities. Therefore, ratings may have been higher due to pain and difficulty with movement.

Table 7-15 Correlation coefficients for actions involving small finger forces and fine manipulations

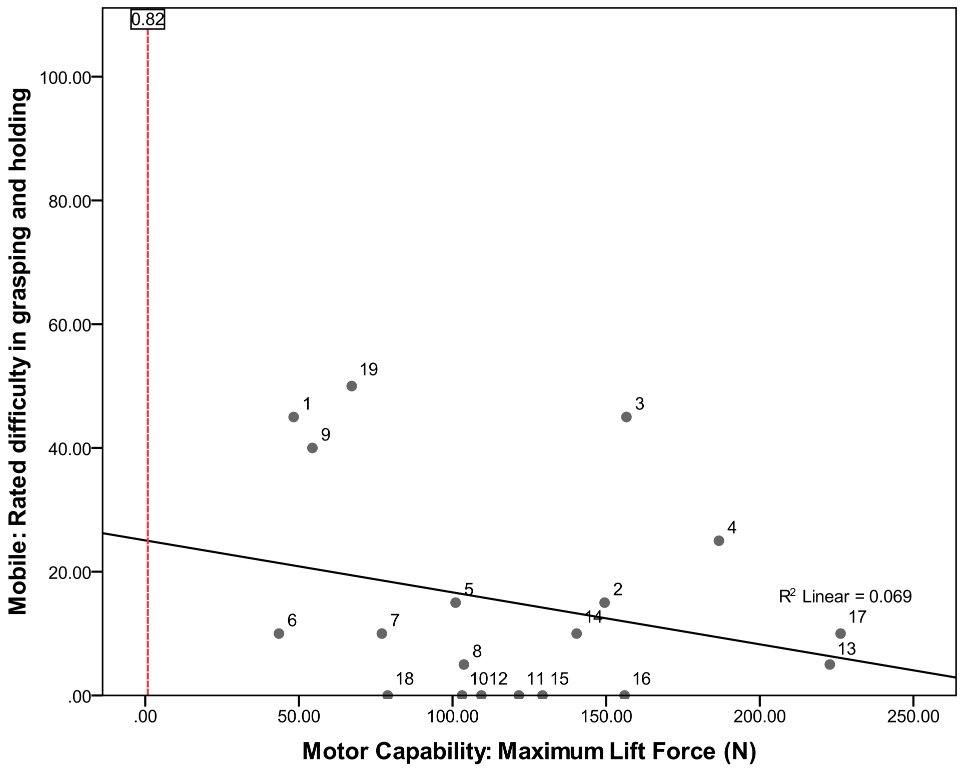
Product	Action	Correlation Coefficient
Clock	Pushing buttons	-0.105
Mobile	Pressing buttons	0.181
Blender	Pushing pulse button	-0.232
Blender	Turning dial	-0.145
Vacuum Cleaner	Sliding power switch	0.017

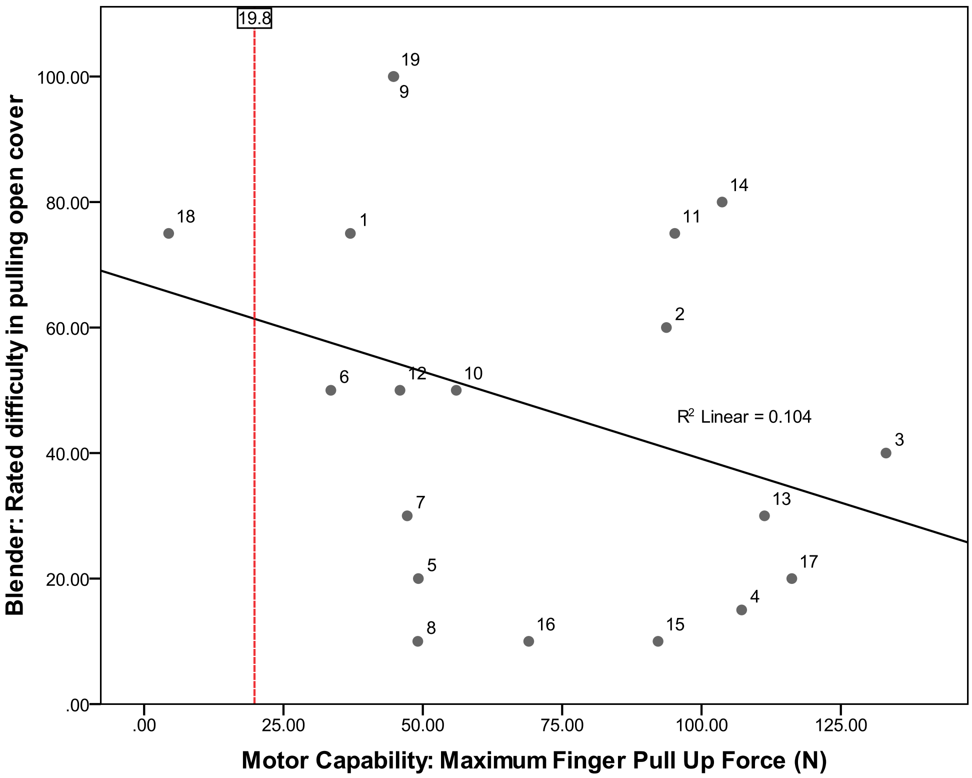
7.3.7.2 Actions Involving Grasping and Large Push/Pull Forces

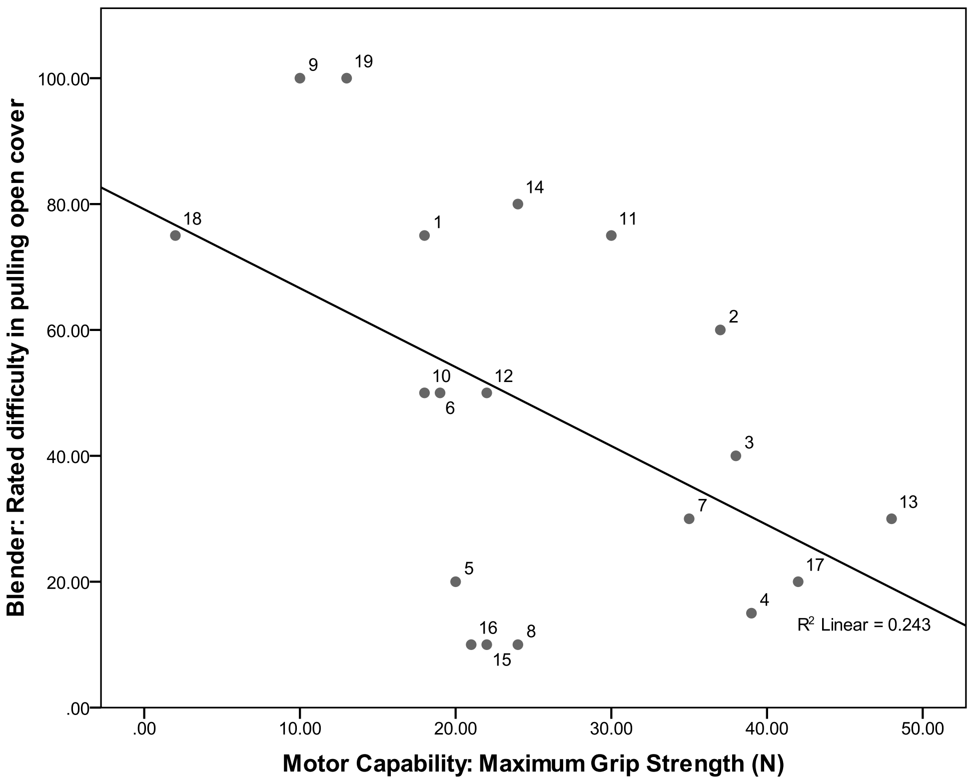
Figure 7-31 shows 6 capability-demand graphs involving grasping with push/pull forces for the mobile and blender. Figure 7-32 shows 3 capability-demand graphs involving grasping with push/pull forces for the vacuum cleaner. Table 7-16 gives the correlation coefficients for all actions. Grasping and holding the mobile r(17)=-0.263 showed a weak linear correlation. Pulling open the blender cover using the maximum finger pull up force measure also resulted in a relatively weak linear correlation r(17)=-0.323. Using maximum grip strength r(17)=-0.492, $p < 0.05$ and maximum lift force r(17)=-0.468, $p < 0.05$ as the capability measures resulted in slightly better linear correlations with rated difficulty. Lifting the blender jug r(16)=-0.042 and pouring from the jug r(16)=-0.022 resulted in poor linear correlations.

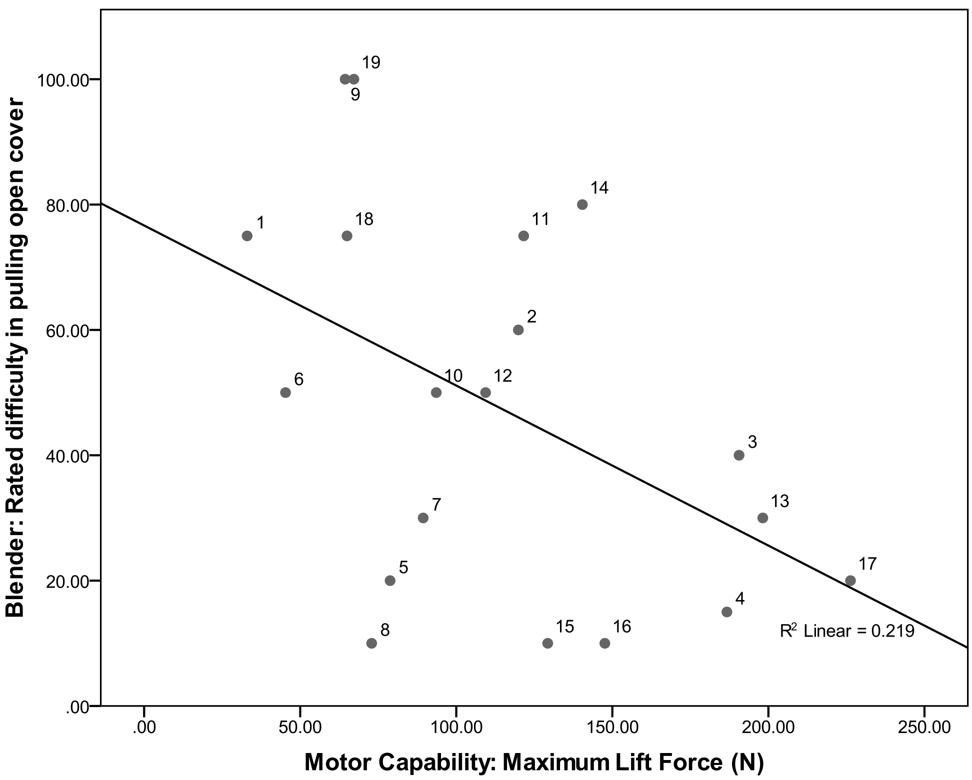
Table 7-16 Correlation coefficients for actions involving grasping and large push/pull forces

Product	Action	Correlation Coefficient
Mobile	Grasping and holding phone	-0.263
Blender	Pulling and open cover (Finger pull up force)	-0.323
Blender	Pulling and open cover (grip strength)	-0.492
Blender	Pulling open cover (maximum lift force)	-0.468
Blender	Lifting jug (maximum lift force)	-0.042
Blender	Lifting and pouring	0.022
Vacuum	Pulling cord out	-0.302
Vacuum	Pushing forward	-0.353
Vacuum	Pulling backward	-0.201


0.82
Mobile: Rated difficulty in grasping and holding
Motor Capability: Maximum Lift Force (N)
R² Linear = 0.069
.00
20.00
40.00
60.00
80.00
100.00
50.00
100.00
150.00
200.00
250.00


19.8
Blender: Rated difficulty in pulling open cover
Motor Capability: Maximum Finger Pull Up Force (N)
R² Linear = 0.104
.00
20.00
40.00
60.00
80.00
100.00
25.00
50.00
75.00
100.00
125.00


Blender: Rated difficulty in pulling open cover
Motor Capability: Maximum Grip Strength (N)
R² Linear = 0.243
.00
20.00
40.00
60.00
80.00
100.00
10.00
30.00
50.00


Blender: Rated difficulty in pulling open cover
Motor Capability: Maximum Lift Force (N)
R² Linear = 0.219
.00
50.00
100.00
150.00
200.00
250.00

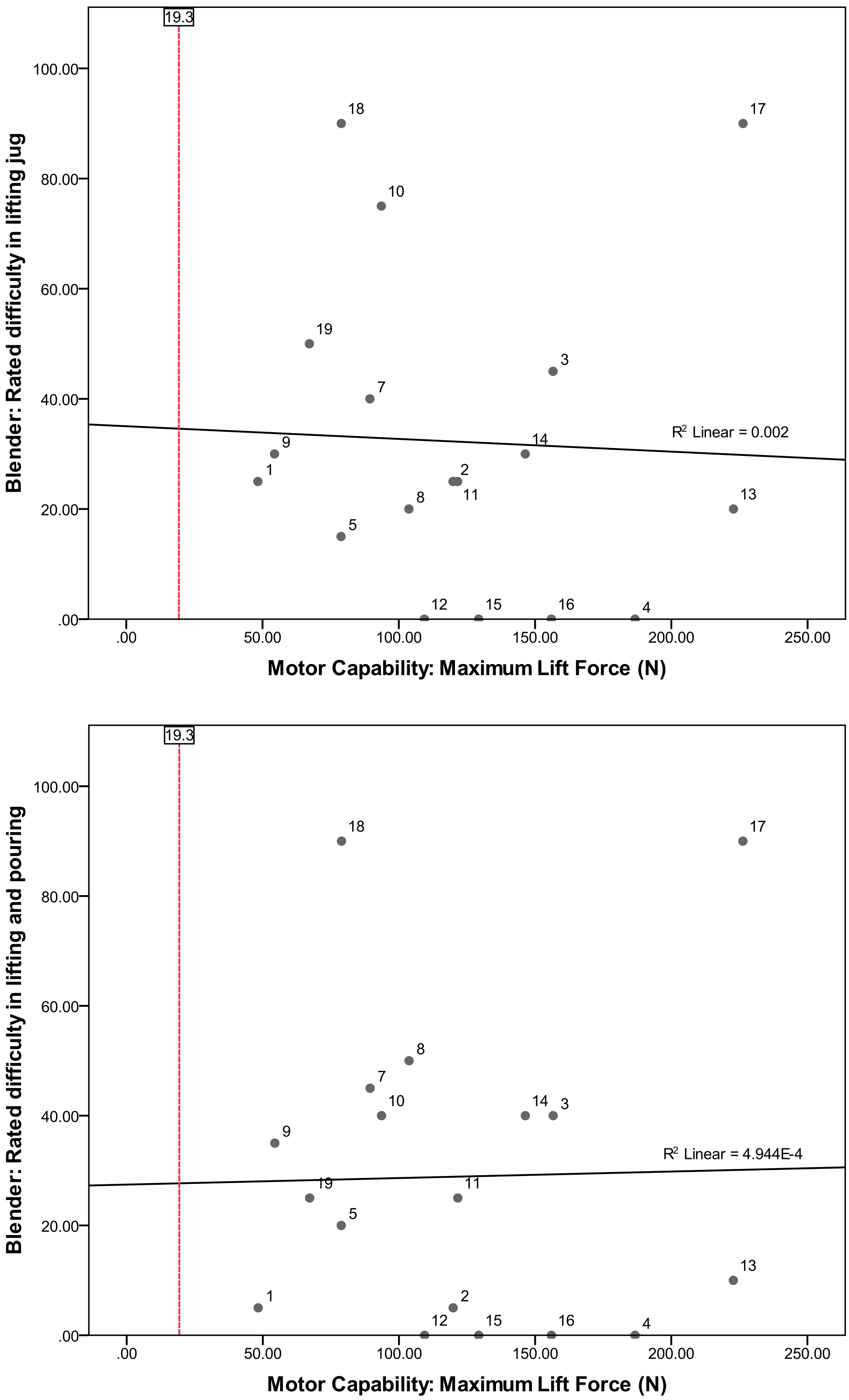
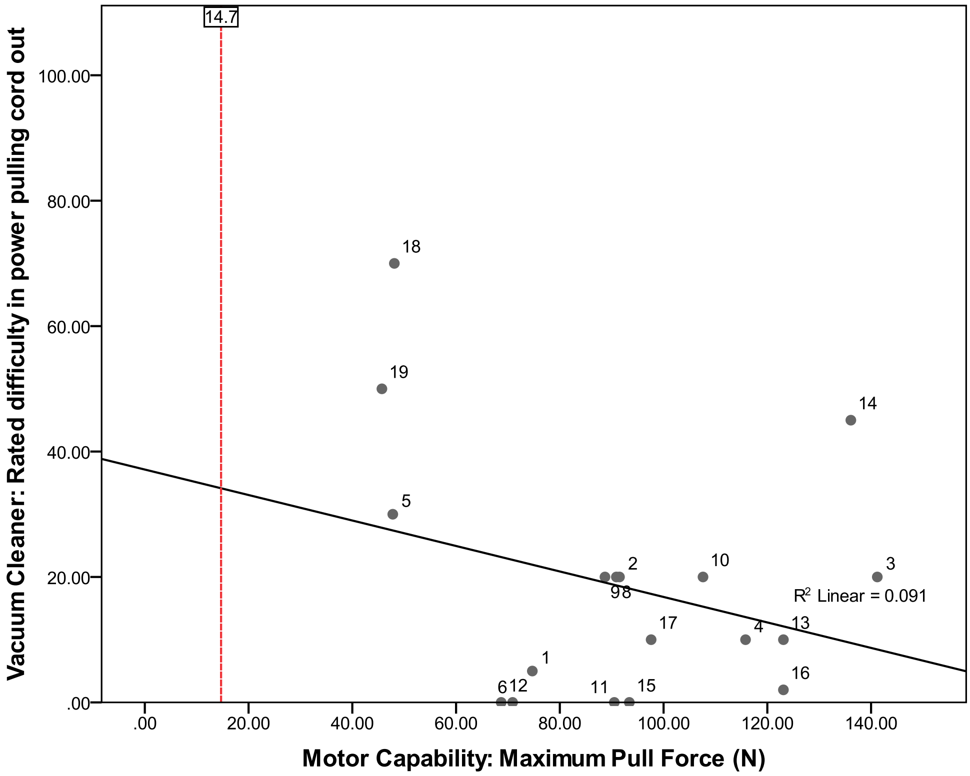
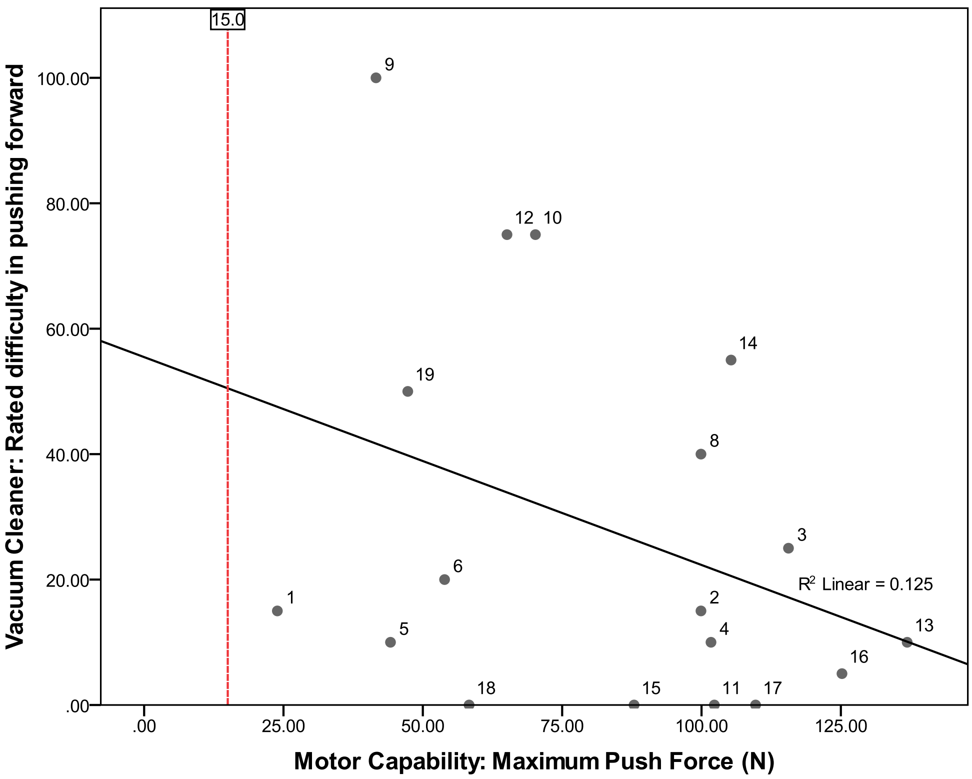


Figure 7-31 Six capability-demand graphs involving grasping with push/pull forces for mobile and blender

Pulling the power cord out r(16)=-0.302, pushing the vacuum cleaner forward r(16)=-0.353 and pulling it backward r(16)=-0.201 all resulted in relatively weak linear correlations (Figure

7-32). Larger push/pull actions such as lifting and opening the blender and moving the vacuum cleaner around showed slightly better linear correlations. Other variables such as balance time and walking speed were weakly correlated with difficulty ratings.





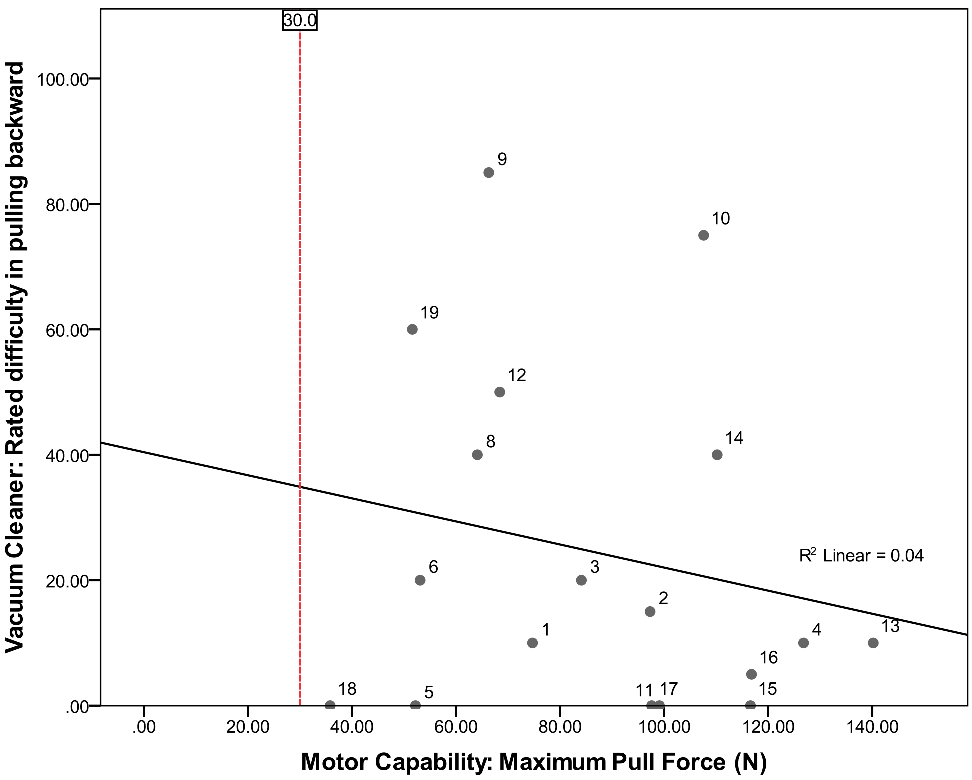


Figure 7-32 Three capability-demand graphs involving grasping with push/pull forces for vacuum cleaner

In general therefore, difficulties experienced with motor actions were weakly correlated with single measurements of maximum force capabilities.

7.3.7.3 Within Product Comparisons

In looking at the motor-capability graphs for each product in turn, there was a common vertical spread of points (similar to the sensory capability-demand graphs), but in addition, there was also a horizontal spread of similar difficulty ratings across the capability scale. This resulted in poor linear relationships for motor actions within the four products. Given that all the motor graphs were of a similar pattern, it can be concluded that there may be other factors at play that influence the ratings and occlude any clear relationships in the data. It is also possible that subjective factors such as pain and frustration in task performance are being factored into the difficulty ratings.

7.3.7.4 Overall Motor Demands

Figure 7-33 shows overall physical demand ratings for each product plotted against the maximum of all physical demand ratings collected for the product. This was plotted to investigate whether participants were making judgments about the overall physical demand of the task based upon the most difficult physical action they had to perform. The correlation coefficients for the clock radio $r(17)=-0.905$, $p< 0.01$, mobile $r(17)=-0.814$, $p< 0.01$, blender $r(17)=-0.693$, $p< 0.01$ and vacuum cleaner $r(16)=-0.663$, $p< 0.01$ all showed a fairly strong positive linear relationship. This indicates that participants were most likely using the most difficult physical action to judge the level of overall physical demand.

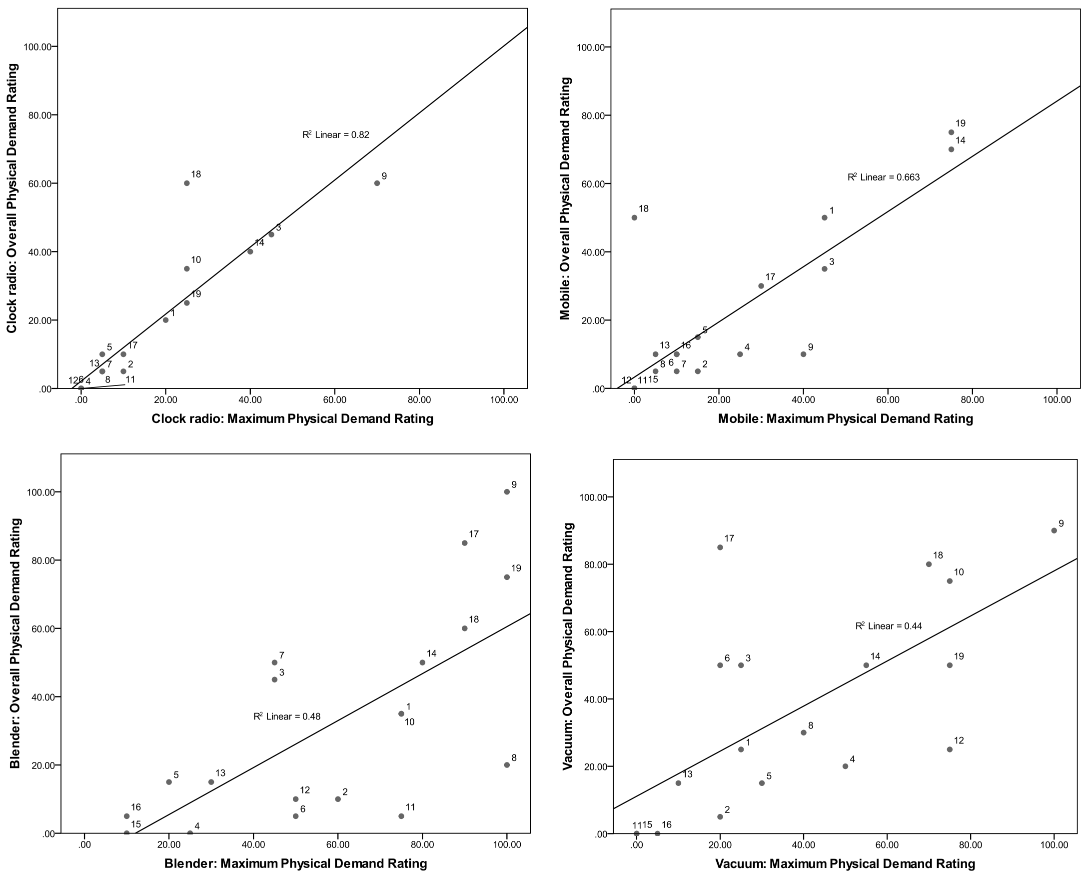


Figure 7-33 Overall physical demand ratings versus the maximum physical demand ratings for each product

However, the correlation coefficients were not as strong for the blender and the vacuum cleaner as they were for the clock radio and the mobile phone. The blender and vacuum cleaner tasks were the more physically challenging tasks of the four tasks, and the graphs show that there was a greater spread of points around the linear fit line. This indicates that there may be other factors at play when participants made the overall difficulty judgment, such as pain, length of the task over which effort is required and possibly even difficulty with other non-motor actions in the task.

7.3.8 Cognitive Capabilities and Demands

In order to investigate the relationships between measured cognitive capabilities and task outcome measures, scatter plots were generated of task time, errors, difficulty starting task, difficulty in selecting subsequent actions and overall mental demand against the four measured cognitive capability variables: (1) verbal working memory (2) visuo-spatial working memory (3) speed of processing and (4) long term memory. These variables appear on the following graphs as ‘DigitSpan’, ‘SSPspanlength’, ‘RTIfivechoicereactiontime’ and ‘GNTpercentcorrect’ respectively.

7.3.8.1 Task time and errors

Table 7-17 gives the correlation coefficients for task time and errors versus the four measured cognitive variables. Scatter plots are included for reference in Appendix 11: Scatter Plots. For task time, the graphs show very weak linear correlations with all four cognitive variables, with no r value being greater than 0.5. The clock radio, blender and vacuum task times show a general decrease with increased span length, while the mobile phone task time showed no clear linear relationship with span length. Task time for the vacuum cleaner shows a general increase with increased reaction time, while the mobile and blender task times show a general decrease with reaction time. The clock radio and mobile task times show a slight general increase with increased long term memory, while the blender and the vacuum cleaner task times show a slight general decrease.

Table 7-17 Correlation coefficients of task times and errors against four cognitive variables

	Digit Span	SSP span length	RTI five choice reaction time	GNT percent correct
Clock radio task time	.198	-.103	-.012	.006
Mobile task time	.015	.030	-.369	.043
Blender task time	.041	-.397	-.207	-.090
Vacuum task time	.187	-.481	.495	-.474
Clock radio errors	-.448	-.341	-.055	-.502
Mobile errors	-.112	.058	-.176	-.059
Blender errors	-.493	-.819	.162	-.638
Vacuum errors	-.516	-.700	.416	-.763

It appears that the general relationship of task times to the cognitive capability variables is different depending on the type of product being considered. It is also likely that for cognitively challenging tasks such as the clock radio task and the mobile phone task, participants with greater capability (e.g. visuo-spatial working memory and long term memory) show longer task times. This could be explained by their period of trial and error behaviour tending to be greater than participants with relatively lower cognitive capability who tend to give up the task quickly if they cannot achieve the end goal.

Errors showed a general decrease with increased digit span for all the tasks, though the strength of the linear correlation was very weak with no r value being greater than -0.516 (vacuum cleaner). The clock radio, blender and vacuum cleaner errors show a general decrease with increased span length, while the mobile phone errors showed no clear linear relationship with span length. Significant linear correlations were found for the blender errors $r(15)=-0.819$, $p< 0.01$ and vacuum cleaner errors $r(14)=-0.700$, $p< 0.01$ with visuo-spatial working memory (Figure 7-34).

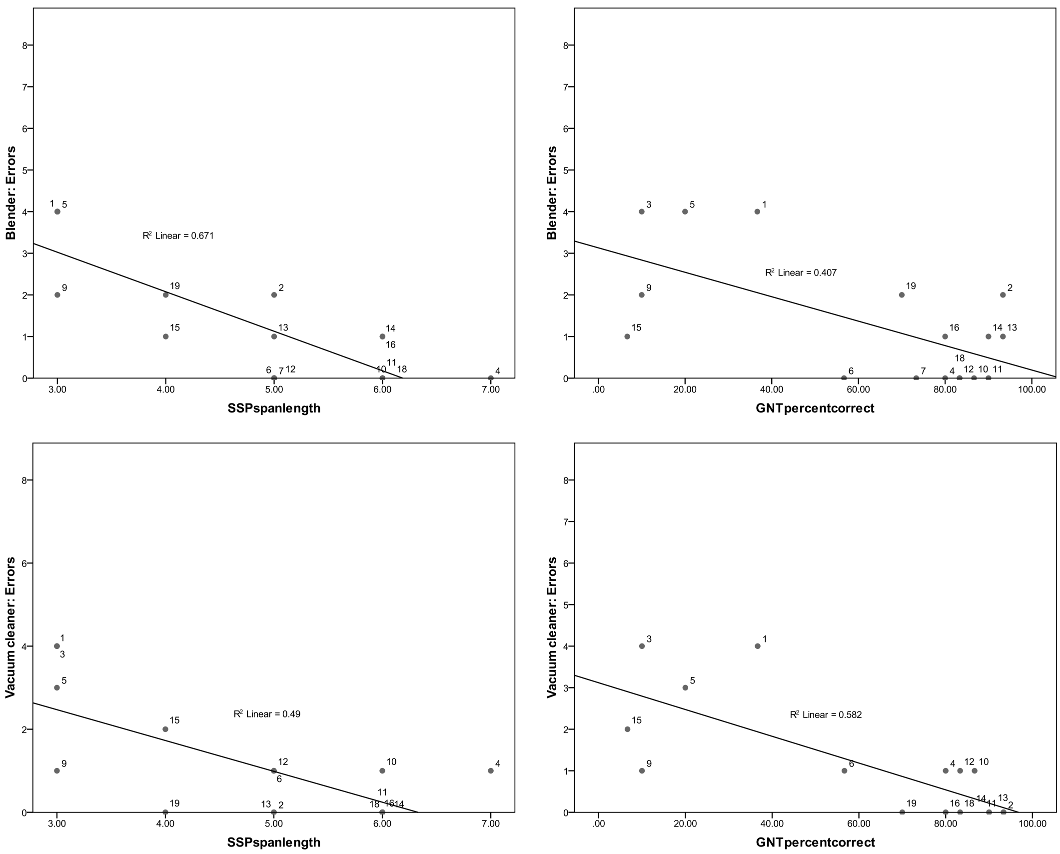


Figure 7-34 Scatter plots of errors vs span length and LTM for the blender and vacuum cleaner

The blender and vacuum cleaner errors show a general increase with reaction time, while the mobile phone errors show a decrease. The clock radio errors show no clear linear relationship with reaction time. The clock radio, blender and vacuum cleaner errors show a general decrease with long term memory, while the mobile phone errors had no significant linear relationship. In particular, blender errors $r(15)=-0.638$, $p< 0.01$ and vacuum cleaner errors $r(15)=-0.763$, $p< 0.01$ show a strong linear relationship with long term memory. Therefore visuo-spatial working memory and long term memory show fairly strong linear relationships with errors for the physically demanding products (blender and vacuum cleaner), but not for cognitively challenging products (clock radio and mobile phone).

7.3.8.2 Starting tasks, next action, overall mental demand and frustration

Table 7-18 gives the correlation coefficients for four difficulty variables (difficulty in figuring out how to start the task, difficulty in figuring out the next action, overall mental demand and task frustration) versus the four measured cognitive variables. The scatter plots are given in Appendix 11: Scatter Plots.

The graphs show very weak linear correlations with all four cognitive variables, with no absolute r value being greater than 0.455 (blender task, span length). The distribution of data points on the graphs show no clear linear relationships, though in most cases the directionality of the fit lines is in keeping with a general decrease in difficulty with an increase in capability. The spread of points on the mobile phone task especially indicates the high levels of difficulty experienced despite the level of cognitive capability.

Table 7-18 Correlation coefficients for cognitive tasks with four capability variables

	Digit Span	SSP span length	RTI five choice reaction time	GNT percent correct
Clock radio: Difficulty starting task	.261	-.212	-.425	-.124
Mobile: Difficulty starting task	.191	.341	-.360	.239
Blender: Difficulty starting task	-.026	-.455	-.260	-.329
Vacuum Cleaner: Difficulty starting task	.190	-.413	.290	-.410
Clock radio: Difficulty with next action	.085	-.266	-.311	-.228
Mobile: Difficulty with next action	.188	.339	-.496	.203
Blender: Difficulty with next action	-.279	-.557	-.165	-.387
Vacuum Cleaner: Difficulty /w next action	.084	-.325	.033	-.377
Clock radio: Overall mental demand	-.089	-.235	-.430	-.241
Mobile: Overall mental demand	.061	.108	-.303	-.006
Blender: Overall mental demand	-.092	-.583	-.223	-.414
Vacuum Cleaner: Overall mental demand	.300	-.284	.304	-.302
Clock radio: Task frustration rating	-.072	-.187	-.022	-.190
Mobile: Task frustration rating	.453	.438	-.511	.521
Blender: Task frustration rating	.134	-.204	-.333	-.148
Vacuum Cleaner: Task frustration rating	.333	.028	.326	-.107

The ‘difficulty with next action’ measure was intended to capture the difficulty (in general) of figuring out what to do to advance toward the end goal. As with starting the task, the graphs show very weak linear correlations with all four cognitive variables, with no absolute r value being greater than 0.557 (blender task, span length). In addition, as with starting the task, the graphs show no clear linear relationships, though in most cases the directionality of the fit lines is in keeping with a general decrease in difficulty with an increase in capability. The mobile phone task graphs go against this trend and indicate very high levels of difficulty experienced despite the level of cognitive capability for most participants (Figure 7-35).

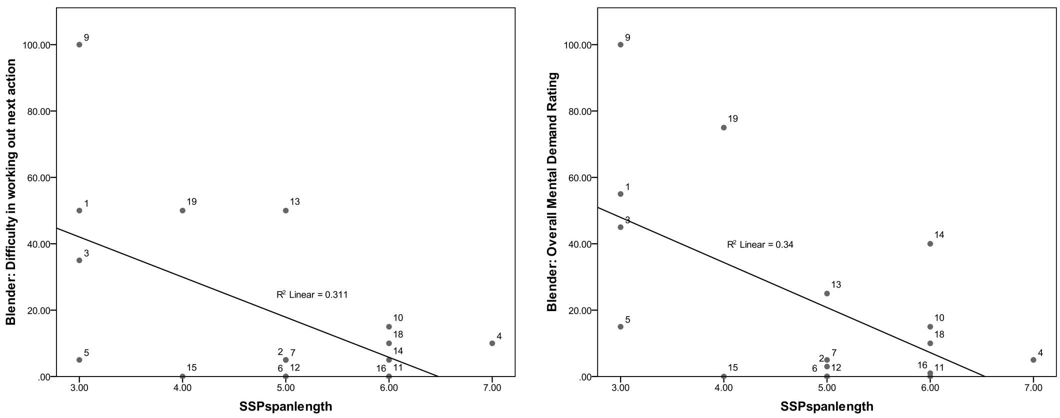


Figure 7-35 Scatter plots of difficulty in working out next action and overall mental demand vs span length for the blender

Overall mental demand shows very weak linear correlations with all four cognitive variables, with no r value being greater than 0.583 (on the blender task related to span length $r(15)=-0.583$, $p< 0.01$). The distribution graphs also show no clear linear relationships, though in most cases the directionality of the fit lines is in keeping with a general decrease in difficulty with an increase in capability. Frustration ratings for the tasks also show weak linear correlations with the four cognitive variables, though there were moderate relationships with all four of the variables for the mobile phone task.

7.3.8.3 Investigating cognitive capability models

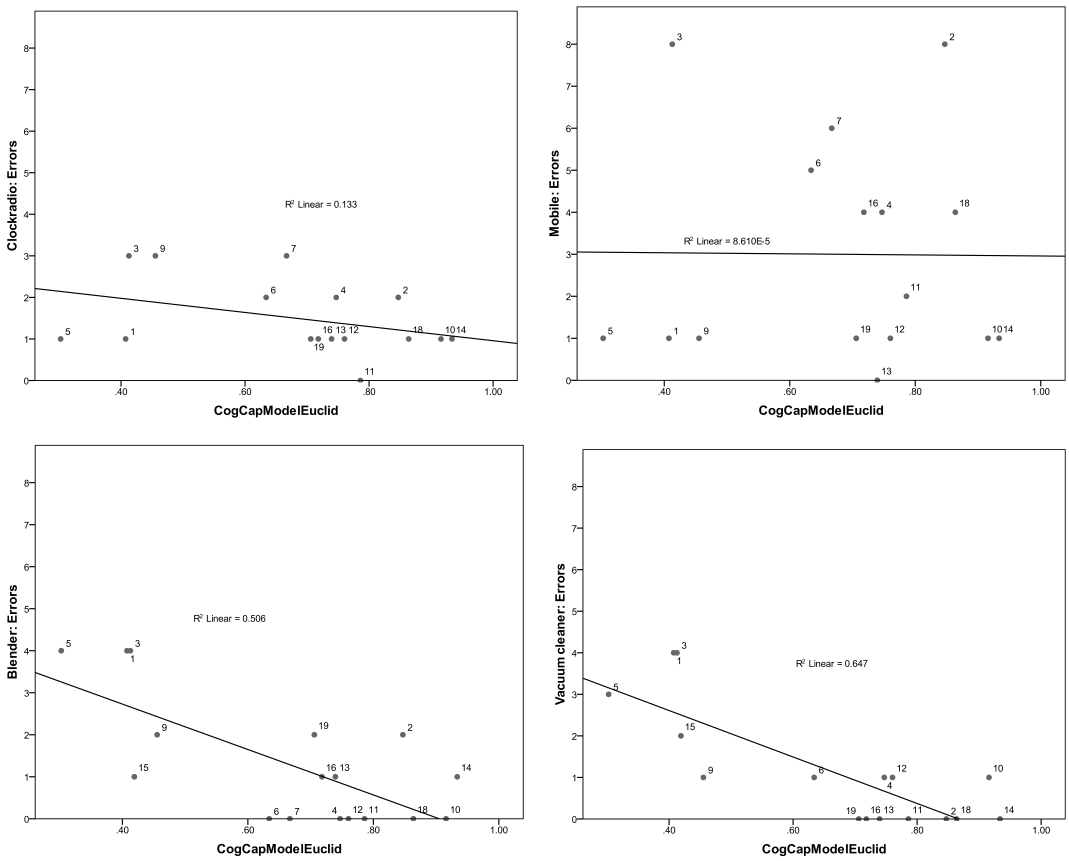
In order to further investigate the relationship between outcome measures and cognitive capabilities, combination models of the four cognitive variables were calculated. These models are described in Table 7-19. The four models used were MAX, MIN, CITY-BLOCK and EUCLIDEAN.

Table 7-19 Combination models for the four cognitive variables (represented by a, b, c and d)

Model	Description	Equation
MAX	The maximum of the four cognitive variables	MAX(a,b,c,d)
MIN	The minimum of the four cognitive variables	MIN(a,b,c,d)
CITY-BLOCK	The city-block metric is the sum of the four cognitive variables	a + b + c + d
EUCLIDEAN	The euclidean metric is the square root of the sum of the squares of the four cognitive variables	$(a^2 + b^2 + c^2 + d^2)^{1/2}$

The four cognitive variables were first scaled to a range between 0 and 1. MAX and MIN models used the maximum and minimum value of the four cognitive variables respectively. The CITY-BLOCK metric used the sum of the four cognitive variables, the result of which was also scaled to a range between 0 and 1. In this case, the CITY-BLOCK measure is equivalent to the arithmetic mean of the cognitive variables. The EUCLIDEAN metric was also used which is the square root of the sum of the squares of the four cognitive variables. MAX and MIN models were used to see the effects of limiting cognitive factors, while the CITY-BLOCK and EUCLIDEAN models (which are both special cases of the Minkowski metric) have been shown to reflect universal principles of generalisation in human cognition (Shepard, 1987). Correlations between these cognitive models and task time, errors, difficulty starting task, difficulty with next action and overall mental demand were investigated.

In general, all cognitive models produced very weak linear relationships with the outcome measures, except for the EUCLIDEAN cognitive capability model in relation to the blender errors and the vacuum cleaner errors. Figure 7-36 shows scatter plots of blender errors $r(15)=-0.711$, $p< 0.01$ and vacuum cleaner errors $r(14)=-0.804$, $p< 0.01$ versus the EUCLIDEAN cognitive capability model. The clock radio and mobile errors, however, showed no clear linear relationship. Figure 7-36 also shows no clear linear relationships between overall mental demand and the EUCLIDEAN cognitive capability model.



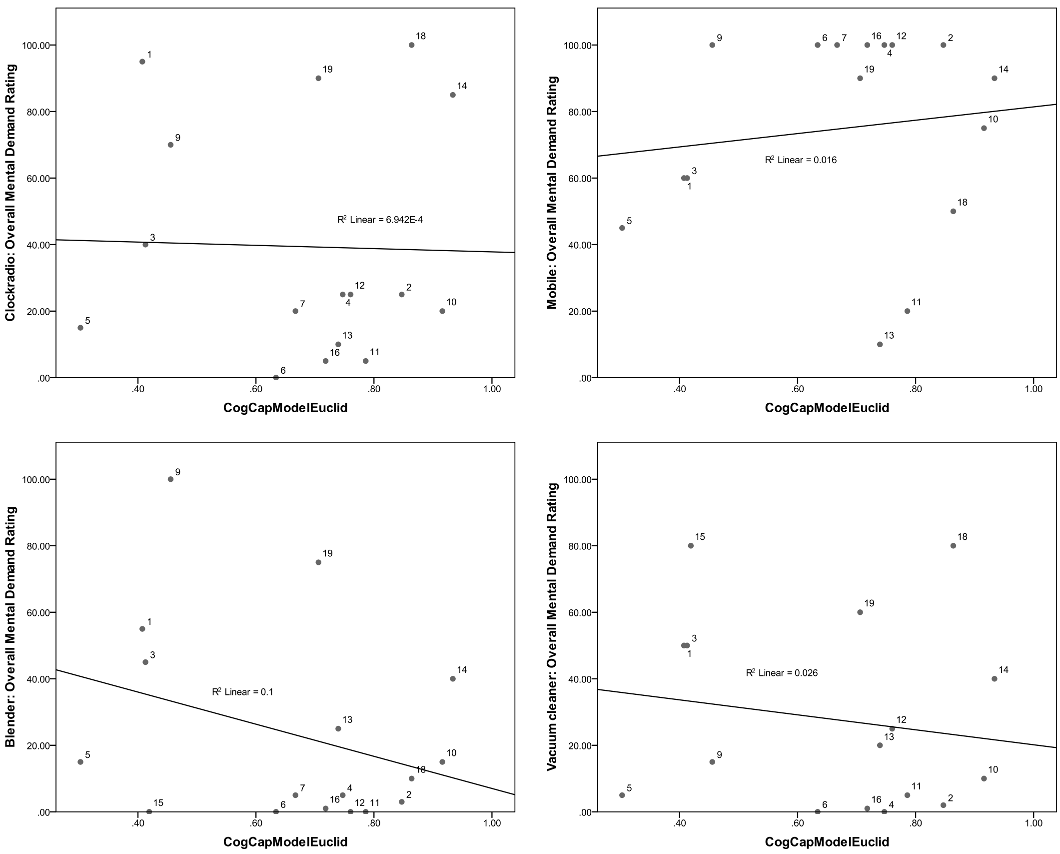


Figure 7-36 Errors and overall mental demand versus a EUCLIDEAN cognitive capability model

7.3.8.4 Relationships with experience

To investigate the relationship of product experience with outcome measures, two measures of product experience were used. At the start of each study session, participants were asked to rate their experience with clock radios, mobile phones, blenders and vacuum cleaners on a scale from 0 to 10. Participants were also asked to describe, in as much detail as they could, the process of using each of these products to perform the tasks used in the study, naming all of the product features they could think of. These descriptions were used to extract the number of steps in the task sequence and the number of features mentioned. These two measures were normalised and combined with equal weighting to give a mental model score for the product category which ranged from 0 to 1. Figure 7-37 shows the relationships between experience ratings and the mental model scores.

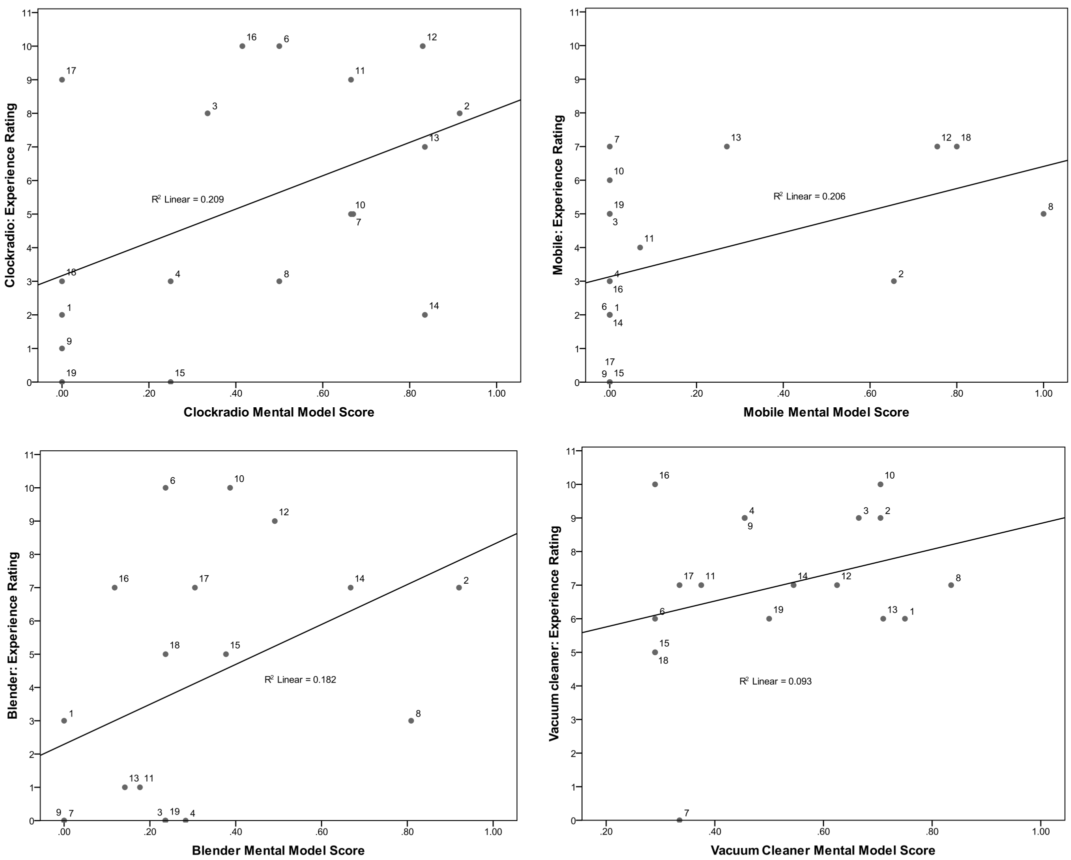


Figure 7-37 Experience ratings versus mental model scores

The two experience measures were only moderately correlated: clock radio $r(16)=0.457$, $p< 0.05$, mobile $r(16)=0.454$, $p< 0.05$, blender $r(16)=0.427$, $p< 0.05$, vacuum cleaner $r(14)=0.305$. The mental model score depended on the capacity of the participant to vocalise their experience utilising their procedural and declarative long term memory. The graphs show that there was significant variation in the study group in terms of matching their self ratings of experience with detailed descriptions of product operation.

Figure 7-38 shows graphs of task times versus experience ratings (with correlation coefficients given in Table 7-20). In general, there was a decrease in task time with an increase in experience rating. Linear relationships were weak for the mobile phone $r(13)=-0.386$, blender $r(16)=-.203$ and vacuum cleaner $r(15)=-0.206$. However, the clock radio task time was moderately correlated with experience ratings $r(13)=-0.609$, $p< 0.01$.

Table 7-20 Correlation coefficients of task times and errors with experience scores

	Experience Rating	Mental Model Score
Clock radio task time	-.609	-.120
Mobile task time	-.386	.066
Blender task time	-.203	.066
Vacuum task time	-.206	-.178
Clock radio errors	-.094	-.095
Mobile errors	.007	.144
Blender errors	-.339	-.211
Vacuum errors	.000	.028

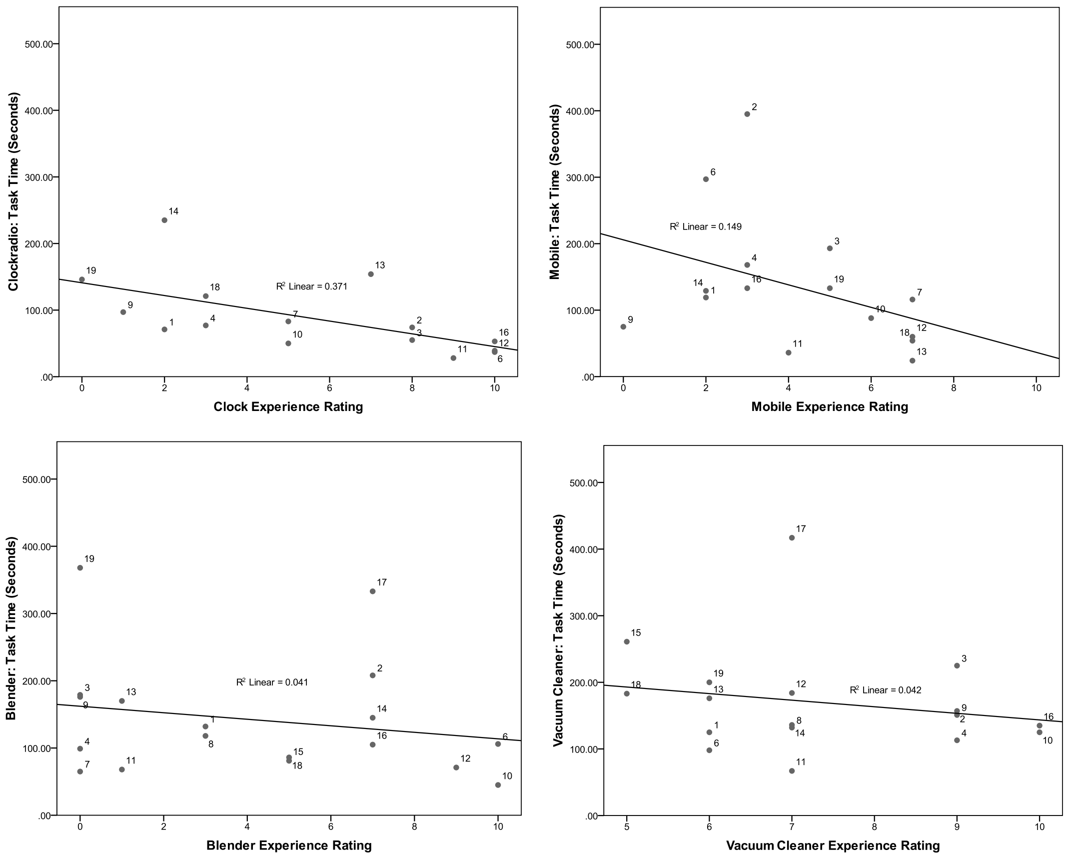


Figure 7-38 Task times versus experience ratings

Errors showed no clear linear relationships with experience ratings, with the best correlation being that for the blender r(16)=-0.339. Correlations of experience ratings with difficulty in starting the task and difficulty in selecting next actions were also fairly weak to moderate. Selecting the next action for the clock radio had the highest correlation of r(14)=-0.557, $p < 0.05$.

Correlations for mental model scores were in general even weaker than experience ratings with respect to task times, errors, difficulty starting the task and difficulty in selecting the next action. The strongest correlations were for difficulty in starting the task r(15)=-0.427, $p < 0.05$ and difficulty in selecting the next action r(15)=-0.562, $p < 0.01$ for the clock radio.

Table 7-21 Correlations of cognitive actions with self rated experience and mental model scores

	Experience Rating	Mental Model Score
Clock radio: Difficulty starting task	-.438	-.427
Mobile: Difficulty starting task	-.439	-.102
Blender: Difficulty starting task	-.369	.135
Vacuum Cleaner: Difficulty starting task	-.436	.149
Clock radio: Difficulty with next action	-.557	-.562
Mobile: Difficulty with next action	-.374	.087
Blender: Difficulty with next action	-.470	-.212
Vacuum Cleaner: Difficulty w/ next action	-.364	.045
Clock radio: Overall mental demand	-.427	-.352
Mobile: Overall mental demand	-.402	-.071
Blender: Overall mental demand	-.401	.035
Vacuum Cleaner: Overall mental demand	-.622	.144
Clock radio: Task frustration rating	-.590	-.253
Mobile: Task frustration rating	-.033	.264
Blender: Task frustration rating	-.145	.036
Vacuum Cleaner: Task frustration rating	-.294	-.082

Figure 7-39 shows the relationships between overall mental demand and experience ratings. The clock radio r(14)=-0.631, $p < 0.01$ and the vacuum cleaner r(15)=-0.622, $p < 0.01$ linear relationships were particularly strong. The mobile phone showed participants with relatively high ratings of mental demand even though they had moderate to high experience ratings, while the blender had participants with low ratings of mental demand even though they had relatively low prior experience ratings. Figure 7-40 shows the relationships between overall mental demand and mental model scores. The correlations for the mobile phone, blender and vacuum cleaner were very weak indicating a lack of any significant linear relationship. The clock radio showed a moderate correlation of r(15)=-0.570, $p < 0.01$.

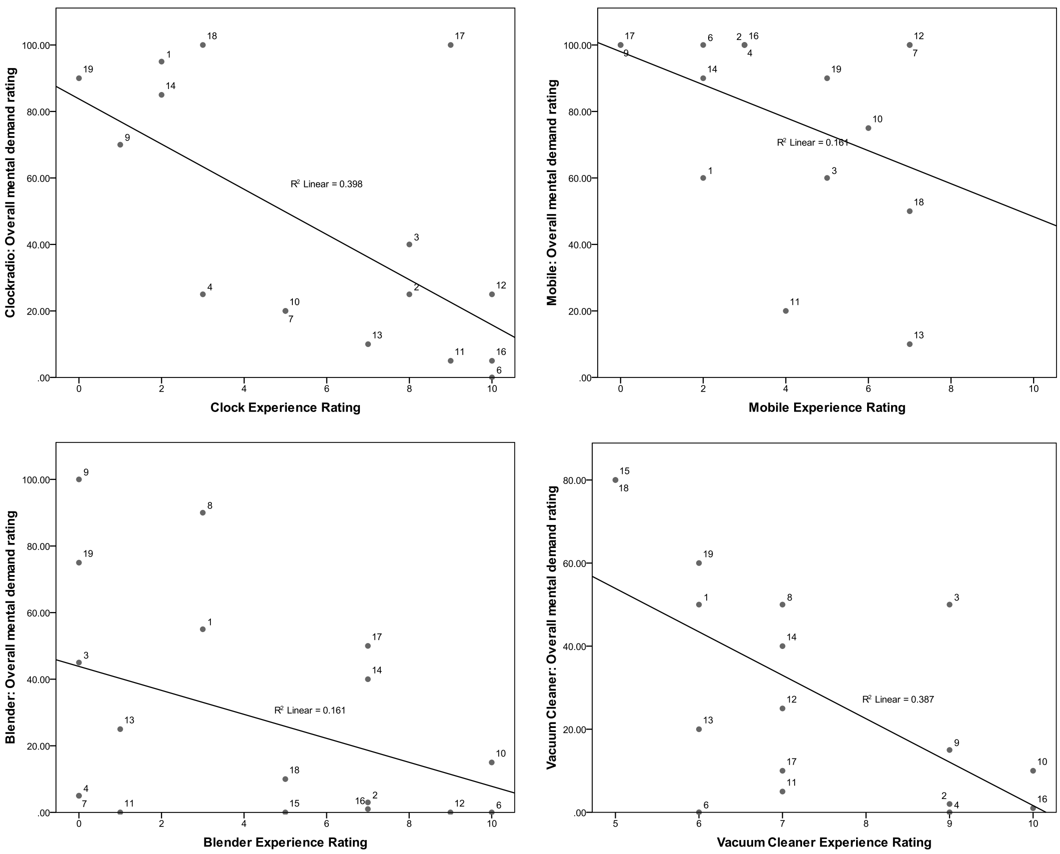


Figure 7-39 Overall mental demand versus experience ratings

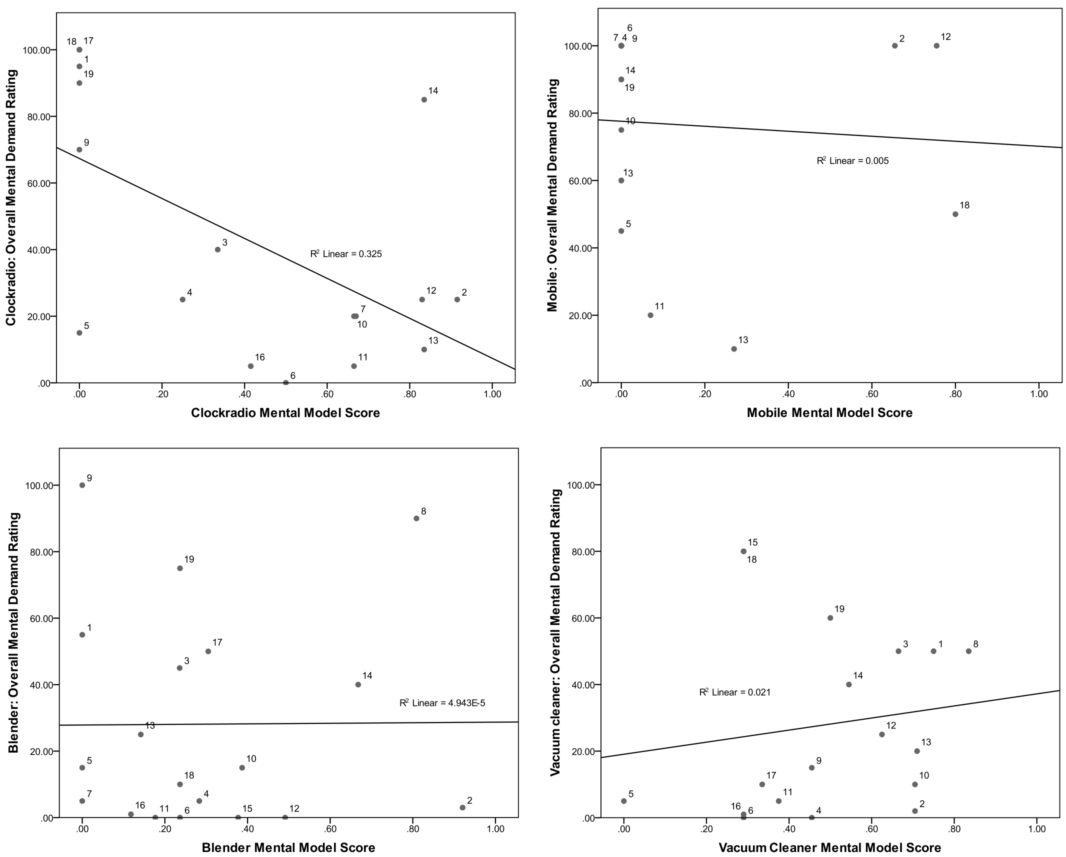


Figure 7-40 Overall mental demand versus mental model scores

7.3.9 Effects of Multiple Variables

Multiple regression can be used as a tool for investigating the relationships between a set of independent or predictor variables and a dependent variable (Dancey & Reidy, 2002; Field, 2009). The method assumes that variables are drawn from a normally distributed population of scores and that they are linearly related to the dependent variable. It also requires in general about 10 to 15 participants per predictor variable used. In addition, multiple regression can be heavily influenced by outliers, and predictors in the model should not be highly correlated with each other i.e. there should be low multicollinearity among predictor variables (Dancey & Reidy, 2002; Field, 2009). There are various methods for constructing linear regression models including hierarchical, forced entry, and stepwise methods (Field, 2009). Field suggests that building models from past research and theoretical considerations is preferred to methods where the computer selects the order of variable entry (stepwise methods).

The sample size in the experimental study was 19 users, and therefore a regression model with at most two predictors could be used for generalisable results. Since the aim of using multiple regression on the data was to investigate which key variables factored in the model (as an indication of which user capabilities came into play in task performance rather than the use of the model itself in prediction), more than two variables were used in model construction. The forced entry method was used for model building by selecting a set of variables that were theoretically related to the outcome measure in question. This resulted in one regression model per analysis. However, the three predictor variables with the greatest standardised beta values in each model were selected to show the importance of predictors in the model (as they indicate the number of standard deviations that the outcome will change as a result of one standard deviation change in the predictor). Therefore these variables will have the most influence on outcome performance measures.

7.3.9.1 Errors and Times Multiple Regressions

Table 7-22 and Table 7-23 show the summarised results of multiple regression models for errors and task times respectively. Independent variables for each model were chosen as the four cognitive capability measures described in previous sections, and full details of each model are given in Appendix 12: Multiple Regression Models. The independent variables in the model are shown ranked from 1st to 3rd in terms of their standardised beta coefficients. The R values (multiple correlation coefficients) for each model are also given.

Table 7-22 Multiple regression models for errors

	Clock radio	Mobile	Blender	Vacuum cleaner
R	0.625	0.351	0.839	0.786
1st	Digit span	Digit span	Span length	Long term memory
2nd	Reaction time	Reaction time	Long term memory	Span length
3rd	Long term memory	Span length	Digit span	Digit span

Table 7-23 Multiple regression models for task times

	Clock radio	Mobile	Blender	Vacuum cleaner
R	0.360	0.502	0.591	0.752
1st	Digit span	Digit span	Span length	Digit span
2nd	Long term memory	Reaction time, Long term memory	Long term memory	Span length
3rd	Reaction time		Reaction time	Long term memory

In Table 7-22, digit span and reaction time are the most important predictors in the clock radio and mobile tasks. For the blender and the vacuum cleaner, span length and long term memory are the most important. This indicates that possibly working memory and overall cognitive processing come into play for cognitively challenging products such as clock radios and mobile phones, whereas products such as the blender and vacuum cleaner rely on long term memory and visuo-spatial capabilities for successful task performance. Table 7-23 shows that digit span is the most important predictor for task times for all products except the blender. Long term memory is the second most important predictor in the models for all products except the vacuum cleaner.

7.3.9.2 Cognitive Multiple Regressions

Table 7-24, Table 7-25 and Table 7-26 show regression models for difficulty in starting the task, difficulty in selecting the next action and overall mental demand. For difficulty in starting the task, reaction time is the most important predictor for all product tasks except the vacuum cleaner. Long term memory is the second most important predictor in the model for the clock radio and the blender, but working memory (span length and digit span) is the second most important predictor for the mobile and vacuum cleaner tasks.

Table 7-24 Multiple regressions for difficulty in starting the task

	Clock radio	Mobile	Blender	Vacuum cleaner
R	0.583	0.475	0.657	0.632
1st	Reaction time	Reaction time	Reaction time	Long term memory
2nd	Long term memory	Span length	Long term memory	Digit span
3rd	Digit span	Digit span	Span length	Span length

Table 7-25 Multiple regressions for difficulty in selecting the next action

	Clock radio	Mobile	Blender	Vacuum cleaner
R	0.502	0.602	0.691	0.659
1st	Reaction time	Reaction time	Span length	Long term memory
2nd	Long term memory	Span length	Reaction time	Digit span
3rd	Span length	Digit span	Digit span	Reaction time

Table 7-26 Multiple regressions for overall mental demand

	Clock radio	Mobile	Blender	Vacuum cleaner
R	0.518	0.394	0.732	0.661
1st	Reaction time	Reaction time	Reaction time	Digit span
2nd	Span length	Span length	Span length	Long term memory
3rd	Long term memory	Long term memory	Long term memory	Span length

For difficulties in selecting the next action, reaction time is the most important predictor for the clock radio and mobile tasks, while span length and long term memory are the most important predictors for the blender and vacuum cleaner tasks respectively. For overall mental demands, reaction time and span length are the most important predictors for the clock radio, mobile and blender tasks. For the vacuum cleaner task however, digit span and long term memory are the most important predictors in the model.

7.4 Discussion of results

7.4.1 Key Findings

The results presented in the previous sections indicate that, given the limitations of the sample size of 19 participants in the study, individual measures of low-level visual (visual acuity, contrast sensitivity), cognitive (digit span, spatial span, reaction time, long term memory), and motor capabilities (grasp forces, push/pull forces) in general correlate weakly to moderately with outcome measures such as task times, errors and rated difficulty.

In the case of vision, most of the correlations for reading text were high, indicating that the essential low-level capabilities necessary for real world task performance were being captured to a large degree. The correlations for seeing product features on the vacuum cleaner were also relatively high, though they were much weaker for the other products. As previously mentioned, the approximation used for seeing product features based on a letter chart test versus a gratings test is a possible explanation for this discrepancy. However, the low-level

measures of visual capability did not correlate well with rated difficulty in all cases, and the individual factors which contribute to this are unclear.
The linear relationships were very weak between rated difficulty and motor capabilities for small finger forces and large push/pull forces with grasping. Most graphs demonstrated a vertical spread of difficulty scores for a given level of capability. As previously mentioned, it is possible that difficulty scores could be influenced by other factors such as discomfort and pain. It was observed that most participants who had higher ratings of difficulty at high capability values were participants who had conditions such as arthritis, deformities and other ailments. Judgments of overall physical demand correlated well with the physical action given the highest difficulty rating in the task. This indicates that participants may be using the most difficult physical action as the benchmark for the overall physical difficulty of the task.

The analysis of measured cognitive capabilities in relation to outcome measures yielded interesting results. Task time showed poor linear relationships with the four cognitive measures for the four products. Errors showed similar poor correlations with the four cognitive measures in the clock radio and mobile phone tasks, but showed strong linear relationships with visuo-spatial working memory and long term memory in the blender and vacuum cleaner tasks. Since these two tasks required spatial movement and orientation for product operation, the results indicate that good visuo-spatial working memory played a significant role in task success. Also, since a generic measure of long term memory correlated well with the blender and vacuum cleaner tasks, it suggests that a general measure of long term memory could be useful in predicting task errors for certain kinds of products. As visuo-spatial working memory and long term memory did not correlate well with errors for the clock radio and mobile phone tasks, it can be inferred that these tasks were not specifically dependent on visuo-spatial reasoning and general knowledge of how products work.

Measures of difficulty for starting the task, working out the next action and overall mental demand were weakly correlated with the cognitive capability measures, with visuo-spatial working memory giving the strongest relationships with these three for the blender task. With the calculation of cognitive capability scores using different models, the EUCLIDEAN model had the best linear correlations with outcome measures, especially with errors for the blender and vacuum cleaner task. Though this suggests that linear models do not account particularly well for the relationships between cognitive capability measures and outcome measures, it may well be the case that other non-linear capability combinations are at work. This is an area that requires further investigation.

The self rated experience measures for the products had stronger relationships with outcome measures than the mental model scores. However, these relationships were weakly correlated overall, though the clock radio had some moderate to fairly strong relationships with task times, errors and overall mental demand.

In looking at the multiple regression models, it is evident that different capabilities come into play depending on the task at hand. The ranking of predictors in the models showed that there is variation in the importance of predictors among different products. The models indicate that certain product tasks may place different types of cognitive loads on users, as in the case of the clock radio and mobile versus the blender and vacuum cleaner for errors made. The R values for the blender and vacuum cleaner were fairly strong while the clock radio and the mobile showed low to moderate R values. This indicates that the multiple regression models showed weak linear dependence between predictors and outcome variables and also performed poorly in accounting for the variance in the data for the clock radio and mobile.

From the results presented, it appears that task demands (the actions required together with the characteristics of the product features) determine the set of underlying user capabilities that are engaged. This combined set of capabilities work together to enable task success. Further, the adaptable nature of the human user allows for variations and substitutions in this set of required capabilities to still enable success in task performance in certain circumstances. A predictive model might therefore utilise a combined capability score based on how the user can compensate for some specific capability loss i.e. a set of common coping strategies or capability substitutions. Another possibility is a model with a built in decision engine that can intelligently combine a generic set of capabilities giving the probabilities of success in specific tasks.

For example, if an individual whose physical capability measures indicated that that he/she would be unlikely to succeed in a particular motor task, the model could check alternative physical actions and force requirements to see if the user could achieve the same goal in another way. In the case of perceptual capability, visual capability loss could be compensated for by multimodal cues such as audio cues or haptics. Cognitive memory impairments could be supported by learnt strategies (meta-cognition) for interaction or by greater reliance on well-established crystallised memory or general reasoning using declarative and general procedural knowledge stored in mental models. Importantly, current data sets and models do not provide for this type of in-built reasoning, but such an approach could produce more robust and realistic results.

In striving to achieve a parsimonious model of user capability that can support Inclusive Design evaluation, the results indicate that alternatives to linear models should also be investigated for describing human capability in disabled populations. It could be that the disabled human user taps into multiple low-level capabilities in non-linear ways, relying on a system of accommodation and coping strategies that would be better modelled with non-linear models. Approaches such as Nonlinear Causal Resource Analysis (NCRA) (Gettman et al., 2003 ; Kondraske et al., 1997) should also be investigated and compared.

7.4.2 Study Limitations

The study aimed to capture the multivariate capability profile of participants in the sensory, cognitive and motor domains, and relate these measures to objective and subjective outcome measures of task performance. In conducting the study, it was observed that participants with low scores on cognitive capability variables were sometimes not able to fully understand instructions. For example, this caused difficulty in getting comfortable and maximum measures of force exertion as some participants did not understand the difference. In addition, even though these participants practised using the visual rating scale a few times before actually using it to rate actions, they misused the scale at times. For example, they would use the anchor points the wrong way around or provide obviously inconsistent ratings. To counter this, if there was any doubt on the part of the researcher, the participants were asked to rate the action again, or the use of the scale was explained again. However, this highlights one operational area of difficulty in reliably measuring the subjective difficulty of people with low cognitive capability.

Participants were asked to make judgements of difficulty on isolated sensory, cognitive and motor actions in a task sequence together with overall judgements of difficulty. In actuality, the human system operates as an integrated whole, and it is difficult to determine the extent to which it is possible to make an individual judgement without being influenced by other actions. Though it was explained to participants that the individual action was to be rated for difficulty, multidimensional judgements would inevitably be difficult to control.

The sample size in the study was relatively small, and the capability measures used were not exhaustive (due to resource and time constraints). Therefore, the results can be considered only indicative, and cannot be generalised broadly to a larger population at this stage. However, the experiment carried out can be considered as a pilot for a larger scale investigation that would be a necessary second step to validate and extend these results. Since such experimental designs are resource intensive, further studies with a larger sample size and adequate coverage of different conditions and disability types would require a major research effort beyond the scope of a single researcher. However, the data collected from such a study would be valuable in determining causal relationships and the characteristics of sub groups that may exist in a heterogeneous group of older and disabled participants. In addition, it would form the basis for investigating different types of mathematical models that could predict task outcomes from capability measures.

7.5 Chapter Summary

The study described in this chapter explored the relationships of low level sensory, cognitive and motor capabilities and outcome measures for 19 older and disabled people interacting with four consumer products. Based on an analysis of linear correlations, the results suggest that in general these relationships are not sufficiently strong to build univariate linear predictive models to support analytical inclusive evaluation. Though in a few cases fairly strong relationships were found, they were not consistent across the four products. It is concluded that further research is required that expands upon the study design in terms of sample size, range of disability covered and utilisation of non-linear modelling and analysis techniques to better understand human capability in older and disabled populations.

Chapter 8 **Conclusions**

The three main research questions outlined in Chapter 3 are revisited in this concluding chapter. Based on the results of the studies conducted, recommendations are made for further work.

8.1 Answering the Research Questions

In Chapter 3, three main exploratory research questions were posed that guided the design and execution of the studies presented in Chapters 4, 5, 6 and 7. Here, each question is revisited with a synopsis of the results from the studies and a discussion of the extent to which each question was successfully answered.

R1: What theoretical models exist for understanding the relationship between human functional capabilities and real world task performance in ageing/disabled populations?

Based on the review presented in Chapter 4, the interface between human capabilities and task demands could be operationalised in two ways: (1) linear combination models of functional capabilities and (2) non-linear resource economic models for the combination of sets of functional capabilities. The first has been used predominantly, as the statistical methods are relatively straightforward and there is precedent for its use in the body ergonomics/human factors literature. The second type of model (non-linear, resource based) has only recently been developed, and experimental studies are ongoing to test its performance.

The first type of model was tested in Chapter 7 in the form of linear regression models of capability variables to model the relationship with task outcome variables. The results demonstrated that a bivariate linear model relating a single capability to a task outcome measure did not adequately account for the variance and shape of the data points (except in the case of reading text in visual tasks). This indicates that human functioning is complex, with multiple capabilities possibly coming into play in task performance. Secondly, the use of multiple regression (multivariate linear models) produced better R values (multiple correlation coefficients) for some products but not others. This may indicate that multivariate linear models could possibly be used as the capability-demand model for some product tasks while alternative models would be required for other product tasks.

A modelling analogy could be drawn in the field of Fluid Dynamics where a fluid undergoes laminar, transitional or turbulent flow depending on certain conditions. The Reynolds number is a value that is calculated to determine, based on the physical conditions, whether the flow is laminar, transitional or turbulent. In a similar way, based on the type of product and task demands, it may be possible to develop metrics similar to the Reynolds number. This metric would indicate if a task could be modelled by multivariate linear models (for the analogous laminar task flow) or some other non-linear model (for the analogous transitional or turbulent task flow). This in essence would be a mixed-model approach. Alternatively, non-linear models may account well for all capability and task domains in disabled populations, but this is subject to further experimental testing and validation.

In summary, the research question R1 was successfully answered by describing the extant theoretical models for understanding the relationship between human functional capabilities and real world task performance in ageing/disabled populations. The work both validated and challenged certain aspects of the models reviewed, which sets the stage for further investigation into predictive model building.

R2: What are the key elements of human functional capability that influence inclusive product interaction?

Chapter 5 presented the main set of low-level capabilities that are required for interacting with a wide range of consumer products. This set was extracted from the body of ergonomics/human factors literature on ageing, disability and Inclusive/Universal Design. The qualitative effect of these capabilities on product interaction was examined in Chapter 6 where it was shown that product design could be improved by taking into account the sensory, cognitive and motor capability losses of older and disabled users.

It was found that these key elements of human functional capability are well described in the literature, both in the field specific literature and in the Ergonomics/Human Factors literature. In a multi-disciplinary and application oriented field such as Inclusive Design, the knowledge base is dependent on the theoretical developments and experimental findings in core fields such as psychology, medicine, engineering, and statistical modelling. Given this situation, theories and models of Inclusive Design would need to integrate new findings in these other fields as they become available. However, there was a consistent set of capabilities that emerged from the literature that was found to account for most interaction problems. Given that the aim is to eventually build a practical model for Inclusive Design evaluation, this set should comprise a fairly comprehensive starting point. Therefore, the research question R2 was successfully answered.

R3: What relationships exist between measures of human functional capability and measures of task performance in the context of a user capability-product demand model of interaction?

Chapter 7 presented an experimental study that examined the relationships between measured sensory, cognitive and motor capabilities and task outcome measures (times, errors, and difficulty). It was found that in general, individual measures of capability are only weakly to moderately linearly related to task outcome measures. It appears that a more realistic model of user capability involves the consideration of the entire set of capabilities required by a particular task.

Further, the modifications and substitutions to this set of capabilities allowing for human adaptability is an area that has not been adequately addressed thus far. These results are indicative at this stage given the limitations of the small sample size and the selected set of capabilities measured. However, the results raise important questions for Inclusive Design evaluation in terms of creating robust and valid capability databases that go beyond utilising single capability measures for task demands. Better models must be found, and alternatives to linear combinations models (for e.g. non-linear combination models) should also be investigated and compared.

An attempt was made to analyse the data collected in light of the ERM model and the Non-Linear Causal Resource Analysis (NCRA) (Kondraske, 2006a; Kondraske et al., 1997). However, this was not possible due to the NCRA requiring a different methodology (the construction of resource demand functions). Given the limitations of the sample size, the reduced set of capabilities measured, and the use of only linear models in describing the data, the research question R3 was partially answered. R3 therefore remains a ripe area for further investigation.

8.2 Recommendations and Further Work

Expanded Studies: Given the limitations of the experimental study presented in Chapter 7, further studies following a similar methodology should be carried out using a larger sample size and a wider range of product interfaces. The sample should contain users from different categories of disability and severity. With a larger data set, relationships between capabilities and outcome measures could be studied with generalisable results. In addition, provisions should be made for investigating non-linear capability combination models in addition to linear combination models in the design of such studies. In this way, adequate data would be provided to further the development of inclusive interaction models that could support design evaluation. There is also scope for the development of models that can incorporate statistical estimates of population parameters such as the time-variation of capabilities within age groups and disability trajectories. Only with the results of such studies could the specification of a user capability database be developed.

Further studies would involve the administration of appropriate sensory, cognitive and motor performance assessments. With large samples, there is a trade-off between cost and the detail of measurement possible. Because of this trade-off, large surveys tend toward subjective assessments while detailed objective measurements can be managed with smaller samples. Since objective capability measurements are desirable to build and test predictive models, there is a need for the development of objective capability tests that can be applied with relative ease. This would allow useful measurements to be taken that are quick and cost-effective to administer but also provide maximum predictive value.

Zachary (2001) outlines three criteria for assessing the validity of a model in engineering decision making: (1) predictive validation; (2) construct validation and, (3) face validation. Predictive validation is the most powerful and entails the comparison of the model to actual performance data. Construct validation is the next best and involves demonstrating that the underlying constructs in the model are valid. Finally, face validation is the least powerful method requiring reviews by experts to assess the apparent validity. However, predictive validation is the most expensive and face validation is the least costly leading to a cost versus power trade-off. The research presented in this dissertation is the first step in this process, and addresses elements of face and construct validation of inclusive assessment models. Predictive validation will need to compare the performance of capability-demand models to user's interaction with real products and examine the accuracy of predictions of quantitative differences in exclusion resulting from design alternatives.

Difficulty Functions: As part of the model building process, the concept of building 'difficulty functions' for different populations could be investigated. A user could be completely excluded from using a product due to the lack of sufficient capability. However, for users with capabilities near this threshold, there would be different levels of difficulty experienced. Though difficulty is a subjective and complex concept, there is value in trying to develop average 'difficulty functions' that could estimate rough levels of difficulty for population sub-groups. Designers could then estimate the numbers of people who would find the task difficult as well as those who would be outright excluded.

Mental Models: There is scope for further development of the mental model concept in analytical evaluation. The study in Chapter 6 utilised a state diagram representation to generate various mental models of the product interface which were used to analyse the product for cognitive load. Firstly, research is needed to investigate how the combination of mental model representations with cognitive workload modelling could be used to predict design exclusion. Secondly, computational mental modelling tools are required that would allow designers to easily generate mental model representations for product interfaces. The software could handle the combinatorial complexity of plotting different paths through the state space of the product. These tools could be linked to cognitive capability data to provide estimates of design exclusion, provided the cognitive predictive models are proven to be valid and robust.

8.3 Reflections on the Research Process

In the process of scientific discovery, it is often useful to take a step back and critically examine established theories, methods and assumptions in order to chart a way forward. This approach was taken in this dissertation when it was found that there is a lack of understanding of the theoretical aspects of Inclusive Design evaluation and a lack of guidance for industry on how to evaluate inclusive product designs (Warburton, 2003). The research project initially started as an effort to develop new evaluation scales for product assessment in Inclusive Design. However, when this fundamental lack of understanding was discovered after reviewing the body of literature, the research focus was re-directed to address this problem.

Successfully working with older and disabled users requires the development of an empathic approach. Patience, understanding and a willingness to understand the individual life context surrounding each study participant are key skills that are needed. In addition, great care is required when communicating with older and disabled people in a research context, as it is necessary to maintain the user's self confidence while at the same time avoiding researcher bias. More preparation time is required for researchers conducting studies with older and disabled users in order to develop some of the skills mentioned above. In addition, lengthier setup and debriefing times are also required to make participants comfortable. Volunteering at charities and becoming involved in social networks are two strategies that could be employed.

8.4 Thesis Contributions

Firstly, existing theories of inclusive interaction were integrated and a generic framework for analytical evaluation was set forth in Chapter 4. This framework serves as a roadmap for further research into designing a system for quantitative predictions of excluded populations in analytical evaluation. The first two requirements were addressed in the framework i.e. the essential areas of human functioning were outlined and the relationships between multiple capability loss and product interaction were investigated.

The analysis of population data on ageing and disability is common in the fields of demography and economics. This data is used primarily for making policy decisions (Grundy et al., 1999). The data from the 1996/97 Great Britain Disability Follow-up Survey (DFS) dataset was used by Keates and Clarkson in a novel way for product assessment given the lack of any other population representative data sets (Keates & Clarkson, 2003a). In order to extract further understanding of the structure of disability in the UK population, this data set was analysed in a new way using exploratory clustering methods (Chapter 5). Though clustering has been used to group disability profiles (Carlsson et al., 2002), it has not been used to look at groupings of disability dimensions. The clustering revealed that disability in the population is structured as a set of related, co-occurring capability losses, and that the simplified 3-axis Inclusive Design Cube masks this real multidimensional structure.

Various studies have looked at the design and evaluation of consumer products with a small sample of users (Cardoso, 2005; Cassim & Dong, 2003; Davis, 1994; Evamy & Roberts, 2004; Kanis, 1993; Langdon et al., 2007). Most studies in the literature look at categories of capability loss in isolation. However, in order to study the issue of multiple capability loss, a small scale qualitative study was designed (Chapter 6) to look at deeper issues arising from the interaction of capability loss, followed by a quantitative study (Chapter 7) that used a novel multiple-capability, multiple-product study design to explore the range of relationships emerging from interaction.

The study in Chapter 6 highlighted important issues for consideration and further research including coping strategies, sensory-cognitive interaction, learning and trial and error behaviour, the trade-off between aesthetics and accessibility, and the context of use. The results also demonstrated that a thorough analytical evaluation can capture some of the important problems that users encounter in real-world interaction.

Most studies focus on a single product family for evaluation at a time (Cardoso, 2005; Langdon et al., 2010; Waller et al., 2009b; Yoxall et al., 2010b), but in reality, people use a single product from a range of different product categories. Another important difference between the study in Chapter 7 and other studies is the investigation of the concept of difficulty as it relates to multiple capability loss and product usage. This takes into account subjective factors in the context of use (Yoxall et al., 2010c) in the form of the perceived level of difficulty in performing various actions with the chosen products. In order to start the model building process described in the generic framework for analytical evaluation, the relationships between capability measures, product demands and task outcome measures were explored using linear relationships. Based on the findings, it was found in general that the linear relationships did adequately not account for the data.

The studies also provided some design relevant guidelines. For example, tasks that require pressing and holding a ‘set’ button while another button is used to cycle through various options should be avoided as much as possible. This type of design sets up an almost insurmountable barrier to the completion of the task. This problem was evident in the design of the memory settings in toaster 2 in Chapter 6 and the design of the clock radio in Chapter 7. In addition, designers should think of inclusive design in terms of trade-offs and multiple capability loss rather than focusing on individual disabilities in isolation. Designers also need to think about how they can design products for use by people with a wide range of coping strategies rather than make assumptions that users will interact is ‘standard’ ways.

Further work is needed in building robust and valid interaction models for use in analytical Inclusive Design evaluation. The ultimate aim is to provide designers with comprehensive analytical assessment methods and user data. This would allow them to make quantitative estimations of design exclusion and difficulty, and thereby develop better products leading to a more inclusive world.

References

Adams, R., Langdon, P., & Clarkson, P. J. (2002). A Systematic Basis for Developing Cognitive Assessment Methods for Assistive Technology. In S. Keates, P. Langdon, P. J. Clarkson & P. Robinson (Eds.), *Universal Access and Assistive Technology* (pp. 53-62): Springer.

Altman, B. M. (2001). Disability Definitions, Models, Classification Schemes, and Applications. In G. L. Albrecht, K. D. Seelman & M. Bury (Eds.), *Handbook of Disability Studies*: Sage Publishers.

Baddeley, A. D. (2000). The Episodic Buffer: A New Component of Working Memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. D. (2002). Is Working Memory Still Working? *European Psychologist, 7*(2), 85-97.

Bajekal, M., Harries, T., Breman, R., & Woodfield, K. (2004). Review of Disability Estimates and Definitions: National Centre For Social Research.

Baldwin, C. L. (2002). Designing in-vehicle technologies for older drivers: application of sensory-cognitive interaction theory. *Theoretical Issues in Ergonomics Science, 3*(4), 307-329.

Bellerby, F., & Davis, G. (2003). *Defining the limits of inclusive design.* Paper presented at the Include 2003 - Inclusive design for society and business, London.

Benktzon, M. (1993). Designing for our future selves: the Swedish experience. *Applied Ergonomics, 24*(1), 19 - 27.

Benyon, D., Crerar, A., & Wilkinson, S. (2001). Individual Differences and Inclusive Design. In C. Stephanidis (Ed.), *User Interfaces For All* (pp. 21-46). London: Lawrence Erlbaum Associates.

Bittner, A. C. (1979). Toward a Computerized Accommodated Percentage Evaluation (CAPE) Model for Automotive Vehicle Interiors. *Society of Automotive Engineers Transactions, 87*, 1285-1299.

Bittner, A. C., & Moroney, W. F. (1975). *The Acommodated Proportion of a Potential User Population: Compilation and Comparisons of Methods for Estimation.* Paper presented at the 18th Annual Meeting of the Human Factors Society, Santa Monica, California.

Bix, L., de la Fuente, J., Pimple, K., & Kou, E. (2009). Is the test of senior friendly/child resistant packaging ethical? *Health Expectations, 12*(4), 430-437.

Bontoft, M., & Pullin, G. (2003). What is an inclusive design process? In J. Clarkson, R. Coleman, S. Keates & C. Lebbon (Eds.), *Inclusive Design - Design for the whole population* (pp. 520-531). London: Springer.

Boren, T., & Ramey, J. (2000). Thinking aloud: reconciling theory and practice *IEEE Transactions on Professional Communication, 43*(3), 261 - 278.

Bridger, R. S. (2003). *Introduction to ergonomics* (Second ed.): Taylor & Francis.

Brinck, T. (2005). Return on Goodwill: Return on Investment for Accessibility. In R. G. Bias & D. J. Mayhew (Eds.), *Cost-Justifying Usability: An Update for the Internet Age* (Second ed.): Morgan Kaufmann Publishers.

BSI. (2005). BS 7000-6:2005 Design management systems - Part 6: Managing inclusive design - Guide: British Standards Institute.
Byrne, M. D. (2003). Cognitive Architecture. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp. 97-117). New Jersey: LEA.
Cambridge Engineering Design Centre. (2011). Inclusive Design Toolkit, 2011, from http://www.inclusivedesigntoolkit.com/
Cañas, J. J., Antolí, A., & Quesada, J. F. (2001). The role of working memory on measuring mental models of physical systems. *Psicologica, 22*, 25-42.
Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*: Lawrence Erlbaum Associates Inc.
Cardoso, C. (2003). Assessment for inclusive design. In J. Clarkson, R. Coleman, S. Keates & C. Lebbon (Eds.), *Inclusive Design - Design for the whole population* (pp. 454-474). London: Springer.
Cardoso, C. (2005). *Design for Inclusivity: Assessing the Accessibility of Everyday Products.* (PhD Thesis), University of Cambridge, Cambridge.
Cardoso, C., & Clarkson, P. J. (2006). *Impairing designers: using calibrated physical restrainers to empathise with users*. Paper presented at the 2nd International Conference for Universal Design, Kyoto, Japan.
Carlsson, G., Iwarsson, S., & Sthål, A. (2002). The Personal Component of Accessibility at Group Level: Exploring the Complexity of Functional Capacity. *Scandinavian Journal of Occupational Therapy, 9*, 100-108.
Carmichael, A. (1999). Style guide for the design of interactive television services for elderly viewers, from http://www.computing.dundee.ac.uk/projects/UTOPIA/publications/Carmichael%20-%20DesignStyleGuideFinal.pdf
Carroll, J. B. (1993). *Human cognitive abilities : a survey of factor-analytic studies*. Cambridge: Cambridge University Press.
Cassim, J. (2004). Cross-market Product and Service Innovation - the DBA Challege Example. In S. Keates, J. Clarkson, P. Langdon & P. Robinson (Eds.), *Designing a More Inclusive World* (pp. 11-19): Springer.
Cassim, J., & Dong, H. (2003). Critical users in design innovation. In J. Clarkson, R. Coleman, S. Keates & C. Lebbon (Eds.), *Inclusive Design - Design for the whole population* (pp. 532-553). London: Springer.
Charness, N., & Bosman, E. A. (1992). Human factors and age. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 495-551). Hillsdale, N.J.: L. Erlbaum Associates.
Charness, N., & Bosman, E. A. (1994). Age-related changes in perceptual and psychomotor performance: Implications for engineering design. *Experimental Aging Research, 20*(1), 45-59.
Clarkson, J., Coleman, R., Keates, S., & Lebbon, C. (Eds.). (2003a). *Inclusive Design - Design for the whole population*. London: Springer.
Clarkson, J., Dong, H., & Keates, S. (2003b). Quantifying design exclusion. In J. Clarkson, R. Coleman, S. Keates & C. Lebbon (Eds.), *Inclusive Design - Design for the whole population* (pp. 422-436). London: Springer.
Clarkson, J., & Keates, S. (2003a). *A framework for minimising design exclusion.* Paper presented at the Include 2003 - Inclusive design for society and business, London.
Clarkson, J., & Keates, S. (2003b). *User capabilities and product demands.* Paper presented at the Include 2003 - Inclusive design for society and business, London.
Clarkson, P. J., & Coleman, R. (2010). Inclusive design. *Journal of Engineering Design, 21*(2-3), 127-129
Clarkson, P. J., Coleman, R., Hosking, I., & Waller, S. (Eds.). (2007). *Inclusive Design Toolkit*. Cambridge: Cambridge Engineering Design Centre.
Coleman, R. (2006). *From margins to mainstream: why inclusive design is better design.* Paper presented at the The 2006 Ergonomics Society Annual Lecture.

Colenbrander, A. (2003). Aspects of vision loss - visual functions and functional vision. *Visual Impairment Research, 5*(3), 115-136.
Cooper, A. (1999). *The inmates are running the asylum : why high-tech products drive us crazy and how to restore the sanity*. Indianapolis, Ind.: Sams.
Crepeau, E. B., Cohn, E. S., & Schell, B. A. B. (2003). *Willard and Spackman's Occupational Therapy* (Tenth ed.): Lippincott Williams & Wilkins.
Crilly, N., Moultrie, J., & Clarkson, P. J. (2004). Seeing things: consumer response to the visual domain in product design. *Design Studies, 25*(6), 547-577.
Crilly, N., Moultrie, J., & Clarkson, P. J. (2009). Shaping things: intended consumer response and the other determinants of product form. *Design Studies, 30*(3), 224-254.
Cutkosky, M. R., & Wright, P. K. (1986). *Modeling manufacturing grips and correlation with the design of robotic hands.* Paper presented at the IEEE International Conference on Robotics and Automation, San Francisco, CA.
Dancey, C., & Reidy, J. (2002). *Statistics without Maths for Psychology. Using SPSS for Windows* (2nd ed.): Pearson Prentice Hall.
Davis, A. (1994). *Hearing in Adults*: Singular Pub Group.
DeBonis, D. A., & Donohue, C. L. (2004). *Survey of audiology : fundamentals for audiologists and health professionals*. Boston, Mass.: Pearson/ Allyn and Bacon.
Dong, H. (2002). *Inclusive design diagnosis based on the ONS scales?* Paper presented at the Cambridge Workshop on Universal Access and Asistive Technology (CWUAAT), Cambridge.
Dong, H. (2004). *Barriers to Inclusive Design in the UK.* (PhD PhD Thesis), University of Cambridge, Cambridge.
Dorst, K. (2003). *Understanding Design - 150 Reflections on Being a Designer*: BIS Publishers.
DTI. (2000). *A study of the difficulties disabled people have when using everyday consumer products*: Department of Trade and Industry
Government Consumer Safety Research.
Edwards, S. J., Buckland, D. J., & McCoy-Powlen, J. (2002). *Developmental & functional hand grasps*. Thorofare, NJ: Slack.
Electronic Industries Alliance, & Electronic Industries Foundation. (1996). Resource Guide for Accessible Design of Consumer Electronics: A Joint Venture of the Electronic Industries Alliance and the Electronic Industries Foundation.
Etchell, L., & Yelding, D. (2004). Inclusive design: products for all consumers. *Consumer Policy Review, 14*(6), 186-193.
Evamy, M., & Roberts, L. (2004). *In Sight: Guide to Design with Low Vision in Mind*: RotoVision.
Everitt, B., Landau, S., & Leese, M. (2001). *Cluster Analysis* (4th Edition ed.). London: Arnold.
Feeney, R. (2002). *Specific anthropometric and strength data for people with dexterity disability*: UK Department of Trade and Industry.
Feeney, R. J., Chrisholm, L., Petherick, N., & Summerskill, S. (1998). *Sampling disabled people for experimental trials.* Paper presented at the Improving the Quality of Life for the Eurpoean Citizen.
Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.): Sage Publications Ltd.
Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2004). *Designing for older adults : Principles and Creative Human Factors Approaches*: CRC Press.
Fisk, J. (1993). Design for the elderly: a biological perspective. *Applied Ergonomics, 24*(1), 47 - 50.
Fleishman, E. A., & Quaintance, M. K. (1984). *Taxonomies of human performance - The description of human tasks*: Academic Press Inc.
Fletcher, D. C., & American Academy of Ophthalmology. (1999). *Low vision rehabilitation: caring for the whole person*. San Francisco ; [Great Britain]: American Academy of Ophthalmology.

Frankfort-Nachmias, C., & Nachmias, D. (2000). *Research Methods in the Social Sciences* (Sixth ed.).
Freudenthal, A. (1999). *The design of home appliances for young and old consumers*. Delft, The Netherlands: Delft University Press.
Fujiura, G. T., & Rutkowski-Kmitta, V. (2001). Counting Disability. In G. L. Albrecht, K. D. Seelman & M. Bury (Eds.), *Handbook of Disability Studies*: Sage Publishers.
Gettman, M., Kondraske, G. V., Traxer, O., Ogan, K., Napper, C., Jones, D. B., . . . Cadeddu, J. A. (2003). Assessment of basic human performance resources predicts operative performance of laparoscopic surgery. *Journal of the American College of Surgeons, 197*(3), 489-496.
Gheerawo, R., & Lebbon, C. (2002). Inclusive design: developing theory through practice. In S. Keates, P. Landgon, J. Clarkson & P. Robinson (Eds.), *Universal Access and Assistive Technology* (pp. 43 - 52). London: Springer-Verlag.
Goodman-Deane, J., Langdon, P. M., & Clarkson, P. J. (2010). Key influences on the user-centred design process. *Journal of Engineering Desig, 21*(2-3), 345-373.
Goodman-Deane, J., Langdon, P. M., Clarkson, P. J., Caldwell, N. H. M., & Sarhan, A. M. (2007). *Equipping designers by simulating the effects of visual and hearing impairments.* Paper presented at the Assets 2007, Tempe, Arizona, USA.
Goodman, J., Dong, H., Langdon, P., & Clarkson, P. J. (2006a). Increasing the uptake of Inclusive Design in industry. *Gerontechnology, 5*(3), 140-149.
Goodman, J., Langdon, P., & Clarkson, P. J. (2007). *Formats for User Data in Inclusive Design.* Paper presented at the HCI International 2007, 4th International Conference on Universal Access in Human-Computer Interaction, Beijing, China.
Goodman, J., Langdon, P. M., & Clarkson, P. J. (2006b). Equipping designers for Inclusive Design. *Gerontechnology, 4*(4), 229-233.
Goonetilleke, T., Case, K., Marshall, R., Porter, M., Sims, R., & Gyi, D. (2003). *Inclusive design by linking ergonomic evaluation and constrain modelling.* Paper presented at the Include 2003 - Inclusive design for society and business, London.
Government Equalities Office. (2010a). Equality Act 2012: What Do I Need To Know? Disability Quick Start Guide. UK: British Chambers of Commerce.
Government Equalities Office. (2010b). Equality Act 2012: What Do I Need To Know? Quick Start Guide for Businesses Who Sell Goods And Services. UK: British Chambers of Commerce.
Grudin, J., & Pruitt, J. (2002). *Personas, Participatory Design and Product Development: An Infrastructure for Engagement.* Paper presented at the PDC 2002.
Grundy, E., Ahlburg, D., Ali, M., Breeze, E., & Sloggett, A. (1999). Disability in Great Britain: Results from the 1996/97 disability follow-up to the family resources survey: UK Department of Social Security.
Gupta, S. P., Keates, S., & Clarkson, P. J. (2003). *Usability considerations in the design of handheld electronic devices.* Paper presented at the International Conferene on Engineering Design ICED 03, Stockholm, Sweden.
Gyi, D. E., Sims, R. E., Porter, J. M., Marshall, R., & Case, K. (2004). Representing older and disabled people in virtual user trials: data collection methods. *Applied Ergonomics, 35*(5), 443 - 451.
Haigh, R. (1993). The ageing process: a challenge for design. *Applied Ergonomics, 24*(1), 9-14.
Hartson, H. R., Andre, T. S., & Williges, R. C. (2001). Criteria for Evaluating Usability Evaluation Methods. *International Journal of Human-Computer Interaction, 13*(4), 373-410.
Hawthorn, D. (2000). Possible implications of aging for interface designers. *Interacting with Computers, 12*(5), 507-528.
Hawthorn, D. (2003). *How universal is good design for older users?* Paper presented at the ACM Conference on Universal Usability, Vancouver, British Columbia, Canada.

Hignett, S., & Wilson, J. R. (2004). The role for qualitative methodology in ergonomics: a case study to explore theoretical issues. *Theoretical Issues in Ergonomics Science, 5*(6), 473-493.

Hitchcock, D. R., Lockyer, S., Cook, S., & Quigley, C. (2001). Third Age Usability and Safety: and Ergonomics Contribution to Design. *International Journal of Human-Computer Studies, 55*(4), 635-643.

Hosking, I. M., Waller, S. D., & Clarkson, P. J. (2010). It is normal to be different: Applying inclusive design in industry. *Interacting with Computers, 22*(6), 496-501.

Hyppönen, H. (1999). *Handbook on Inclusive Design of Telematics Applications*.

Jacko, J. A., Dixon, M. A., Jr, R. H. R., Scott, I. U., & Pappas, C. J. (1999). *Visual profiles: a critical component of universal access.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, Pittsburgh, Pennsylvania, United States.

Jacko, J. A., & Vitense, H. S. (2001). A review and reappraisal of information technologies within a conceptual framework for individuals with disabilities. *Universal Access in the Information Society, 1*(1), 56-79.

John, B. E., & Kieras, D. E. (1994). The GOMS Family of Analysis Techniques: Tools for Design and Evaluation, Human-Computer Interaction Institute Technical Report CMU-HCII-94-106: School of Computer Science, Carnegie Mellon University.

Johnson, D., Clarkson, P. J., & Huppert, F. (2010). Capability measurement for inclusive design. *Journal of Engineering Design, 21*(2-3), 275-288.

Kanis, H. (1993). Operation of Controls on Consumer Products by Physically Impaired Users. *Human Factors, 35*(2), 3.5-328.

Kanis, H. (2003). *Research in Design: situated individuality versus summative analysis.* Paper presented at the IEA 2003.

Kanis, H., & Arisz, H. J. (2000). How many participants: A simple means for concurrent monitoring. *Proceedings of the IEA 2000/HFES 2000 Congress,* 6 - 637.

Karwowski, W. (2002). Achieving compatibility in human-computer interface design and evaluation. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*: LEA.

Keates, S., & Clarkson, J. (2003a). *Countering design exclusion - An introduction to inclusive design*: Springer.

Keates, S., & Clarkson, J. (2003b). Design exclusion. In J. Clarkson, R. Coleman, S. Keates & C. Lebbon (Eds.), *Inclusive Design - Design for the whole population* (pp. 88-102). London: Springer.

Keates, S., & Clarkson, J. (2003c). *Supporting the adoption of inclusive design practices.* Paper presented at the Include 2003 - Inclusive design for society and business, London.

Kieras, D. (2003). Model-based Evaluation. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp. 1139-1164). New Jersey: LEA.

Kisner, C., & Colby, L. (2007). *Therapeutic Exercise: Foundations and Techniques* (Fifth ed.): F.A. Davis Company.

Kondraske, G. V. (1988). *Workplace Design: An Elemental Resource Approach to Task Analysis and Human Performance Measurements.* Paper presented at the Choice For All. ICAART'88. RESNA International Conference of the Association for the Advancement of Rehabilitation Technology, Montreal, Canada.

Kondraske, G. V. (2006a). The Elemental Resource Model for Human Performance. In J. D. Bronzino (Ed.), *Biomedical Engineering Fundamentals (The Biomedical Engineering Handbook)* (3rd ed., pp. 75-01 - 75-19): CRC Press.

Kondraske, G. V. (2006b). Human performance engineering design and analysis tools. In J. D. Bronzino (Ed.), *Biomedical Engineering Fundamentals (The Biomedical Engineering Handbook)* (3rd ed., pp. 84-01 - 84-16): CRC Press.

Kondraske, G. V. (2006c). Human performance engineering: Challenges and prospects for the future. In J. D. Bronzino (Ed.), *Biomedical Engineering Fundamentals (The Biomedical Engineering Handbook)* (3rd ed., pp. 85-01, 85-10): CRC Press.

Kondraske, G. V. (2006d). Measurement tools and processes in rehabilitation engineering. In J. D. Bronzino (Ed.), *Biomedical Engineering Fundamentals (The Biomedical Engineering Handbook)* (3rd ed., pp. 73-01 - 73-16): CRC Press.

Kondraske, G. V., Johnston, C., Pearson, A., & Tarbox, L. (1997). *Performance prediction and limiting resource identification with nonlinear causal resource analysis*. Paper presented at the Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL , USA

Krishnan, V., & Ulrich, K. T. (2001). Product Development Decisions: A Review of the Literature. *Management Science, 47*(1), 1-21.

Kroemer, K. H. E. (1986). Coupling the Hand with the Handle: An improved Notation of Touch, Grip, and Grasp. *Human Factors, 28*(3), 337-339.

Langdon, P., Japikes, R., Clarkson, P. J., & Wallace, K. (2003). *The Effectiveness Of Psychological Distance In An Empirical Study Of Guideline Clustering For Aerospace Design.* Paper presented at the International Conference on Engineering Design ICED 03, Stockholm, Sweden.

Langdon, P., Lewis, T., & Clarkson, P. J. (2007). The effects of prior experience on the use of consumer products. *Universal Access in the Information Society, 6*(2), 179-191.

Langdon, P. M., Lewis, T., & Clarkson, P. J. (2010). Prior experience in the use of domestic product interfaces. *Universal Access in the Information Society, 9*(3), 209-225

Leonard, V. K., Jacko, J. A., Yi, J. S., & Sainfort, F. (2006). Human Factors and Ergonomics Methods. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (3rd Edition ed., pp. 292-321): Wiley.

Marshall, R., Case, K., Porter, M., Summerskill, S., Gyi, D., Davis, P., & Sims, R. (2010). HADRIAN: a virtual approach to design for all. *Journal of Engineering Design, 21*(Issue 2-3), 253-273.

Martin, J., & Elliott, D. (1992). Creating an Overall Measure of Severity of Disability for the Office of Population Censuses and Surveys Disability Survey. *Journal of the Royal Statistical Society. Series A (Statstics in Society), 155*(1), 121-140.

Martin, J., Meltzer, H., & Elliot, D. (1988). OPCS surveys of disability in Great Britain Report 1 - The prevalence of disability among adults.

Mayhew, D. J. (1999). *The Usability Engineering Lifecycle*: Morgan Kaufmann Publishers.

Meister, D. (1985). *Behavioural analysis and measurement methods*. New York ; Chichester: Wiley.

Meister, D. (1995). Simulation and Modelling. In J. R. Wilson & E. N. Corlett (Eds.), *Evaluation of human work : a practical ergonomics methodology* (2nd ed., pp. 202-228). London: Taylor & Francis.

Monk, A. (1998). Cyclic interaction: a unitary approach to intention, action and the environment. *Cognition, 68*, 95-110.

Moore, B. C. J. (1998). *Cochlear hearing loss*. London: Whurr.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). San Diego, CA ; London: Academic Press.

Morris, J. M. (1994). User interface design for older adults. *Interacting with Computers, 6*(4), 373-393.

Moschis, G. P. (1992). Gerontographics: a Scientific Approach to Analyzing and Targeting the Mature Market. *Journal of Services Marketing, 6*(3), 17-26.

Moschis, G. P. (2003). Marketing to older adults: an updated overview of present knowledge and practice. *Journal of Consumer Marketing, 20*(6), 516-525.

MRC. (2008). The Medical Research Council Cognitive Function and Ageing Study (MRC CFAS) from http://www.cfas.ac.uk/

Mueller, J. (2003). Universal products in the US. In J. Clarkson, R. Coleman, S. Keates & C. Lebbon (Eds.), *Inclusive Design - Design for the whole population* (pp. 318-335). London: Springer.

Napier, J. R. (1956). The Prehensile Movements of the Human Hand. *The Journal of Bone and Joint Surgery, 38B*(4), 902-913.

NATCEN. (2008). English Longitudinal Study of Ageing.

Newell, A. F., & Gregor, P. (2000). *User Sensitive Inclusive Design.* Paper presented at the Proceedings on the 2000 conference on Universal Usability.

Newell, A. F., & Gregor, P. (2002). Design for older and disabled people - where do we go from here? *Universal Access in the Information Society, 2*(1), 3 - 7.

Nielsen, J. (1993). *Usability Engineering*. California: Academic Press.

Norman, D. (2005a). Human-Centered Design Considered harmful. *Interactions, 12*(4), 14-19.

Norman, D. A. (1983). Some Observations on Mental Models. In D. Gentner & A. L. Stevens (Eds.), *Mental Models* (pp. 7-14): Lawrence Erlbaum Associates.

Norman, D. A. (2002). *The Design of Everyday Things*. New York: Basic Books (Perseus).

Norman, D. A. (2005b). HCD harmful? A Clarification. Retrieved from http://www.jnd.org/dn.mss/hcd_harmful_a_clarification.html

Norris, B., & Wilson, J. R. (1997). *Designing safety into products - Making ergonomic evaluation part of your design process*. Nottingham: Product Safety and Testing Group, Institute for Occupational Ergonomics, Department of Manufacturing Engineering and Operations Management, University of Nottingham.

Norton, T. T., Corliss, D. A., & Bailey, J. E. (2002). *The Psychophysical Measurement of Visual Function*: Butterworth Heinemann.

Office for Disability Issues. (2011). Disability prevalence factsheet: Disability prevalence estimates 2009/10. Retrieved from http://odi.dwp.gov.uk/disability-statistics-and-research/disability-facts-and-figures.php

Office for National Statistics. (2011). Statistical Bulletin: 2010-based national population projections - principal projection and key variants. Retrieved from http://www.ons.gov.uk/ons/dcp171778_235886.pdf

Park, D. C. (1999). The basic mechanisms accounting for age-related decline in cognitive function. In D. C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer* (pp. 3-21): Psychology Press.

Payne, S. J. (2003). Users' Mental Models: The Very Ideas. In J. M. Carroll (Ed.), *HCI Models, THeories and Frameworks* (pp. 135-156): Morgan Kaufmann Publishers.

Peebles, L., & Norris, B. (1998). *Adultdata: the handbook of adult anthropometric and strength measurements*. London: Department of Trade and Industry, Government Consumer Safety Research.

Persad, U., & Clarkson, P. J. (2005). *Characterising Users for Inclusive Design - Three Requirements.* Paper presented at the Include 2005, Royal College of Art, London, UK.

Persad, U., Langdon, P. M., & Clarkson, P. J. (2005). *Assessing user capabilities from small scale samples.* Paper presented at the HCI International 2005, Las Vegas, Nevada.

Petrie, H. (2001). Accessibility and usability requirements for ICTs for disabled and elderly people: a functional classification approach. In C. Nicolle & J. Abascal (Eds.), *Inclusive Design Guidelines for HCI*: Taylor & Francis.

Pheasant, S. (1987). *Ergonomics - Standards and Guidelines for Designers*. Milton Keynes: British Standards Institution.

Pirkl, J. J. (1994). *Transgenerational Design - Products for an Aging Population*. New York: Van Nostrand Reinhold.

Pirkl, J. J., & Babic, A. L. (1988a). *Guidelines and Strategies for Designing Transgenerational Products: A Resource Manual for Industrial Design Professionals*: Copley Publishing Group.

Pirkl, J. J., & Babic, A. L. (1988b). *Guidelines and Strategies for Designing Transgenerational Products: An Instructor's Manual*: Copley Publishing Group.

Porter, J. M., Case, K., Marshall, R., Gyi, D., & Oliver, R. S. n. (2004). 'Beyond Jack and Jill': designing for individuals using HADRIAN. *International Journal of Industrial Ergonomics, 33*(3), 249-264.

Porter, M., marshall, R., Sims, R., Gyi, D., & Case, K. (2003). *HADRIAN: a human modelling CAD tool to promote 'design for all'.* Paper presented at the Include 2003 - Inclusive design for society and business, London.

Poulson, D., Ashby, M., & Richardson, S. (Eds.). (1996). *USERfit A practical handbook on user-centred design for assistive technology*. Loughborough: HUSAT Research Institute.

Preiser, W. F. E., & Ostroff, E. (Eds.). (2001). *Universal Design Handbook*: McGraw-Hill.

Proctor, R. W., & Vu, K.-P. L. (2003). Human Information Processing: An Overview for Human-Computer Interaction. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp. 35-51). New Jersey: LEA.

RNIB. (1991). Blind and partially sighted adults in Britain: The RNIB survey, from http://www.rnib.org.uk/xpedio/groups/public/documents/visugate/public_survviadu.hcsp

Robson, C. (2002). *Real World Research* (Second Edition ed.): Blackwell Publishing.

Rosson, M. B., & Carroll, J. M. (2002). *Usability Engineering: Scenario-based development of human-computer interaction*: Morgan Kaufmann Publishers.

Rubin, J., & Chisnell, D. (2008). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests* (Second ed.). New York: Wiley.

Sanders, M. S., & McCormick, E. J. (1993). *Human Factors in Engineering and Design* (Seventh ed.): McGraw-Hill Publishers.

Schieber, F. (1992). Aging and the senses. In J. E. Birren, R. Sloan & G. Cohen (Eds.), *Handbook of mental health and aging* (pp. 251-306). New York: Academic Press.

Schiffman, H. R. (2000). *Sensation and perception : an integrated approach* (5th ed.). New York Chichester: Wiley.

Sears, A., & Young, M. (2003). Physical Disabilities and Computing Technologies: An Analysis of Impairments. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp. 482-503). New Jersey: LEA.

Semmence, J., Gault, S., Hussain, M., Hall, P., Stanborough, J., & Pickering, E. (1998). Family Resources Survey, Great Britain 1996-97. London, UK: Department of Social Security, Corporate Document Services.

Sesto, M., Vanderheiden, G., & Radwin, R. G. (2004). *Functional Characterization of Disability and Interface Use.* Paper presented at the RESNA 2004 Conference.

Shepard, R. N. (1987). Toward a Universal Law of Generalization for Psychological Science. *Science, 237*(4820), 1317-1323.

Sims, R. E. (2003). *'Design for All': methods and data to support designers.* (PhD), Loughborough University, Loughborough.

Smith, S., Norris, B., & Peebles, L. (2000). *Older Adultdata: The Handbook of Measurements and Capabilities of the Older Adult*: DTI.

Snyder, R. (2002). Market segmentation: successfully targeting the mature population. *The Journal on Active Aging, March/April 2002*.

Stanton, N., Hedge, A., Brookhuis, K., Salas, E., & Hendrick, H. (2004). *Handbook of Human Factors and Ergonomics Methods*: CRC Press.

Stanton, N. A., & Baber, C. (2002). Error by design: methods for predicting device usability. *Design Studies, 23*(4), 363-384.

Stanton, N. A., Salmon, P. M., walker, G. H., Baber, C., & Jenkins, D. P. (2005). *Human Factors Methods - A Practical Guide for Engineering and Design*: Ashgate Publishing Limited.

Stary, C. (2000). *A Structured Contextual Approach to Design for All.* Paper presented at the 6th ERCIM Workshop "User Interfaces for All", Florence, Italy.

Steenbekkers, L. P. A., & VanBeijsterveldt, C. E. M. (Eds.). (1998). *Design-relevant characteristics of ageing users*. Delft, The Netherlands: Delft University Press.

Steenbekkers, L. P. A., VanBeijsterveldt, C. E. M., Dirken, J. M., Houtkamp, J. J., Molenbroek, J. F. M., & Voorbij, A. I. M. (1999). Design-relevant ergonomic data on Dutch elderly. *International Journal for Consumer & Product Safety, 6*(3), 99-115.
Story, M. F. (2001). Principles of Universal Design. In W. F. E. Preiser & E. Ostroff (Eds.), *Universal Design Handbook*: McGraw-Hill.
Story, M. F., & Mueller, J. L. (2001). Universal Design of Products. In W. F. E. Preiser & E. Ostroff (Eds.), *Universal Design Handbook*: McGraw-Hill.
Story, M. F., Muller, J. L., & Mace, R. L. (1998). *The Universal Design File - Designing for People of All Ages and Abilities*: The Center for Universal Design.
Thimbleby, H. (2007). *Press on: Principles of interaction programming*: MIT Press.
Tibble, M. (2004). User's guide to disability estimates and definitions, from http://www.dwp.gov.uk/asd/
Tsai, W.-C., & Lee, C.-F. (2009). *Hand Requirements of the Interaction Performance in Older Adults Using Small Electronic Products*. Paper presented at the IASDR 2009.
U.S. Department of Justice. (2011). ADA Update: A Primer for Small Business: U.S. Department of Justice, Civil Rights Division, Disability Rights Section.
United Nations. (2002). World Population Ageing 1950-2050. New York: United Nations.
United Nations. (2007). *World Population Ageing 2007*. New York: United Nations Department of Economic and Social Affairs.
United Nations. (2009). *Population Ageing and Development 2009*. New York: United Nations Department of Economic and Social Affairs.
Van der Veer, G. C., & Melguizo, M. d. C. P. (2003). Mental Models. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (pp. 52-80). New Jersey: LEA.
Van der Vegte, W. F. (2002). Consideration and modelling of use processes in computer-aided conceptual design : a state of the art review. *Transactions of the SDPS - Journal of Integrated Design & Process Science, 6*(2), 25-59.
Vanderheiden, G. (2001). Fundamentals and Priorities for Design of Information and Telecommunication Technologies. In W. F. E. Preiser & E. Ostroff (Eds.), *Universal Design Handbook*: McGraw-Hill.
Vanderheiden, G., & Tobias, J. (1999). Barriers, incentives and facilitators for adoption of universal design practices by consumer product manufacturers Retrieved 24th June 2007, from http://trace.wisc.edu/docs/hfes98_barriers/barriers_incentives_facilitators.htm
Vanderheiden, G., & Tobias, J. (2000, Jul 29-Aug 4). *Universal design of consumer products: current industry practice and perceptions.* Paper presented at the Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Association, 'Ergonomics for the New Millennnium', San Diego, CA.
Vanderheiden, G. C., & Vanderheiden, K. (1992). Guidelines for the Design of Consumer Products to Increase Their Accessibility to People with Disabilities - Working Draft 1.7: Trace R & D Center.
VisionConection. (2006). Causes of Vision Impairment Retrieved 4th December 2006, from http://www.visionconnection.org/Content/Research/EpidemiologyandStatistics/Statistics/CausesofVisionImpairment/default.htm
Waller, S. D., Goodman-Deane, J., Langdon, P. M., Johnson, D., & Clarkson, P. J. (2009a). *Developing a method for assessing product inclusivity*. Paper presented at the 17th International Conference on Engineering Design, Stanford, CA, USA.
Waller, S. D., Langdon, P. M., & Clarkson, P. J. (2009b). *Visualizing design exclusion predicted by disability data: A mobile phone case study*. Paper presented at the 5th International Conference on Universal Access in Human-Computer Interaction, San Diego, California, USA.
Waller, S. D., Langdon, P. M., & Clarkson, P. J. (2010a). Designing a more inclusive world. *Journal of Integrated Care, 18*(4), 19-25.

Waller, S. D., Langdon, P. M., & Clarkson, P. J. (2010b). Using disability data to estimate design exclusion. *Universal Access in the Information Society, 9*(3), 195-207.

Warburton, N. (2003). Everyday Inclusive design. In J. Clarkson, R. Coleman, S. Keates & C. Lebbon (Eds.), *Inclusive Design - Design for the whole population* (pp. 250-269). London: Springer.

Warburton, N. (2005). *Mainstream inclusive success.* Paper presented at the Include 2005 International conference on inclusive design, Royal College of Art, London, UK.

Wickens, C. D., & Hollands, J. G. (1999). *Engineering Psychology and Human Performance* (3rd ed.): Prentice Hall.

Wilcox, S. (2007). *Effective Communication of User Needs to Product Designers*. Paper presented at the Include 2007: Designing with People, London.

Winter, M., & Lemke, M. (2005). MED-AUDIT Impairment Categories: Working towards Mapping AMI Usability *Resource Document*: Rehabilitation Engineering Research Center on Accessible Medical Instrumentation.

Woodson, W. E., Tillman, B., & Tillman, P. (1992). *Human factors design handbook : information and guidelines for the design of systems, facilities, equipment, and products for human use* (2nd ed.). New York ; London: McGraw-Hill.

World Health Organisation. (1992). *International Statistical Classification of Diseases and Related Health Problems (Tenth Revision)* (Vol. 1): World Health Organization (WHO).

World Health Organisation. (2001). *International Classification of Functioning, Disability and Health*: World Health Organization (WHO).

World Health Organisation. (2011). World Report on Disability: World Health Organization (WHO).

Yoxall, A., Janson, R., Bradbury, S. R., Langley, J., Wearn, J., & Hayes, S. (2006). Openability: Producing Design Limits for Consumer Packaging. *Packaging Technology and Science, 19*(4), 219–225.

Yoxall, A., Kamat, S. R., Langley, J., & Rowson, J. (2010a). Squeezability. Part 2: getting stuff out of a bottle. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 224*(6), 1261-1271.

Yoxall, A., Langley, J., Janson, R., Lewis, R., Wearn, J., Hayes, S. A., & Bix, L. (2010b). How Wide Do You Want the Jar?: The Effect on Diameter for Ease of Opening for Wide-mouth Closures. *Packaging Technology and Science, 23*(1), 11-18.

Yoxall, A., Langley, J., Musslewhite, C., Rodriguez-Falcon, E. M., & Rowson, J. (2010c). Husband, Daughter, Son and Postman, Hot-water, Knife and Towel: Assistive Strategies for Jar Opening. In P. M. Langdon, P. J. Clarkson & P. Robinson (Eds.), *Designing Inclusive Interactions - Inclusive Interactions Between People and Products in Their Contexts of Use* (pp. 187-196): Springer.

Zachary, W., Campbell, G. E., Laughery, K. R., Glenn, F., & Cannon-Bowers, J. A. (2001). The application of human modeling technology to the design, evaluation and operation of complex systems. In E. Salas (Ed.), *Advances in Human Performance and Cognitive Engineering Research* (Vol. 1, pp. 199-247): Elsevier Science Ltd.

Zajicek, M. (2004). *A methodology for interface design for older adults.* Paper presented at the Proceedings of the Sixth International Conference on Enterprise Information Systems ICEIS 2004, Oporto.

Appendices

Appendix 1: Consent Form for Toaster Study

UNIVERSITY OF CAMBRIDGE
Department of Engineering

edc cambridge

PARTICIPANT NO # _____

Statement of Informed Consent

Purpose

You have been asked to participate in a pilot observational study for assessing the ease of use of two toasters. By participating in this activity, you will help us to provide information to designers to design products that are easier to use. This activity is meant to help us understand how people interact with products; it is not intended to test your individual performance in any way.

Study Procedure

You will be asked to perform simple tasks with the toasters such as making toast and toasting a bagel. While you perform these tasks, I will videotape your interactions and record your comments.

Confidentiality

We will use the data you give us (along with the information we collect from other participants) to develop design guidance and measurement scales for product design evaluation. To ensure confidentiality, we will not associate your name with your data. However, it may be possible to be identified from image and video data.

Freedom to Withdraw

You may withdraw from the activity at any time without penalty.

Data Usage

Data collected in the form of pictures and video may be used in published papers, reports and presentations. This data would be made anonymous (names would not be associated with images and video). If you would prefer that this data not be used, please check the appropriate box below.

☐ Yes, image and video data can be used in papers, reports and presentations

☐ No, please do not use image and video data in papers, reports and presentations

If you agree to these terms, please indicate your acceptance by signing below:

Signature ______________________

Printed Name ______________________

Date ____ | ____ | ____
DAY MONTH YEAR

Appendix 2: Participant Questionnaire for Toasters Study

Toasters Study | Participant Questionnaire

Participant Number: ______

A. General

Name: ____

Age: ________

Gender: □ Male □ Female

Highest Education Level: _______________________________

Current Occupation: ___________________________________

B. Health

1. Do you have any long term health problems or complaints that affect your everyday activities? If so, pease list any current health or medical conditions:

____

____

5) Do you have problems with any of the following?

(a) Your vision

(b) Your hearing

(c) Remembering what you were doing (your memory)

(d) Walking and moving about

(e) Lifting objects such as kettles and pots

(f) Twisting and turning dials and knobs on cookers, microwaves or other products

(g) Grasping, pulling or pushing product handles such as microwave or washing machine doors

If so, please describe the nature of these problems:

____

____

____

6) Do you currently use any aids to help with seeing, hearing or moving about (such as magnifiers, hearing aids or a wheelchair)?

If YES, please list the aids that you currently own or use:

C. Product Experience

1. Do you currently own and use a toaster? □ Yes □ No

If yes, please describe the make/model of the toaster/toasters:

____

____

2. How many years experience do you have with toasters (please check one)?

□ No experience/Never Used a Toaster

□ Less that one year

□ Less than 5 years

□ More than 5 years

3. What does it mean for a product (such as a toaster) to be simple? Please list as many qualities of simple products that you can think of:

____

____

If you have any examples of simple products that you own and use, please list them as well:

____

____

Appendix 3: Ethical Approval Letter

Karen Douglas
Secretary

UNIVERSITY OF CAMBRIDGE

CAMBRIDGE
PSYCHOLOGY RESEARCH
ETHICS COMMITTEE

Mr U Persad
Cambridge Engineering Design Centre
Department of Engineering
Trumpington Street
Cambridge
CB2 1PZ

8 May 2007

Application No: 2007.21

Dear Mr Persad

User Characterisation for Inclusive Product Assessment

The Cambridge Psychology Research Ethics Committee has given ethical approval to your research project: User Characterisation for Inclusive Product Assessment, as set out in your application dated 29 March 2007.

The Committee attaches certain standard conditions to all ethical approvals. These are:

(a) that if the staff conducting the research should change, any new staff should read the application submitted to the Committee for ethical approval and this letter (and any subsequent letter concerning this application for ethical approval);
(b) that if the procedures used in the research project should change or the project itself should be changed, you should consider whether it is necessary to submit a further application for any modified or additional procedures to be approved;
(c) that if the employment or departmental affiliation of the staff should change, you should notify us of that fact.

Members of the Committee also ask that you inform them should you encounter any unexpected ethical issues.

If you will let me know that you are able to accept these conditions, I will record that you have been given ethical approval.

Yours sincerely

K S Douglas

Cc: Prof. P John Clarkson, Dr P Langdon

17 Mill Lane
Cambridge CB2 1RX
Telephone: 01223 766894
Fax: 01223 332355
E-mail: mb422@admin.cam.ac.uk

Appendix 4: Study Consent Form

PARTICIPANT NO # _____

UNIVERSITY OF CAMBRIDGE
Department of Engineering

edc cambridge

Statement of Informed Consent

Purpose

You have volunteered to participate in a study with the aim of assessing the ease of use of some household consumer products. By participating in this activity, you will help us to understand the difficulty experienced when using everyday products. It will also help us to advise designers on how to design products that are easier to use. This study is meant to help us understand how well **the product** matches your capability, and it is **not** a test of your individual performance in any way.

Study Procedure

This study will take approximately two hours. This includes planned breaks. However, you may take a break at any time. Audio and video recording will be used at different points to record information.

1) You will first be asked some general background questions including questions about your current health status, questions about problems you experience in daily life, and questions about your experience with products that you may own and use. This information will be stored anonymously.

2) Next, you will be asked to perform some tasks that will allow me to better understand your capability. These would be assessments of your vision, hearing, memory, and strength. These assessments are **not** meant to test you in any way, but rather to allow us to see if products best match up to your capability.

3) Finally, you will be asked to perform tasks with four consumer products. While you perform these tasks, I will videotape you using the product. After performing each task, I will ask you to rate your difficulty with specific actions.

Use of resulting study data

We will use the recorded data (along with the information we collect from other participants) to develop design guidance which will help designers in evaluating their designs. Data collected in the form of pictures and video may be used in published papers, reports and presentations at conferences and written up in journals. If you would not like this image and video data to be used, you can indicate this when giving your signed consent.

Ethical approval and confidentiality

This study has gained ethical approval from the Cambridge Psychology Research Ethics Committee. To ensure confidentiality, we will not associate your name with your data and consent forms will be kept in a locked file. Recorded data will be referenced by number codes and your name would not be used. Results are normally summarised and presented in terms of groups of study participants. If data

PARTICIPANT NO # _____

about any individual is presented, your name will not be used to identify you. However, it may still be possible to be identified from images and video clips.

Freedom to Withdraw

You may withdraw from the study at any stage without explanation.

Remuneration

You will be offered a £10 Boots voucher for your participation.

Signed Consent

Please sign below to give your written consent and acknowledge the receipt of a £10 Boots voucher:

Can image and video data can be used in papers, reports and presentations?

☐ *YES* - image and video data can be used in papers, reports and presentations

☐ *NO* - please do not use image and video data in papers, reports and presentations

If you agree to these terms, please indicate your acceptance by signing below:

Signature ____________________

Print Name ____________________

Date ______ | ______ | ______
Day Month Year

☐ *Please check this box if you would like a summary of the study results.*

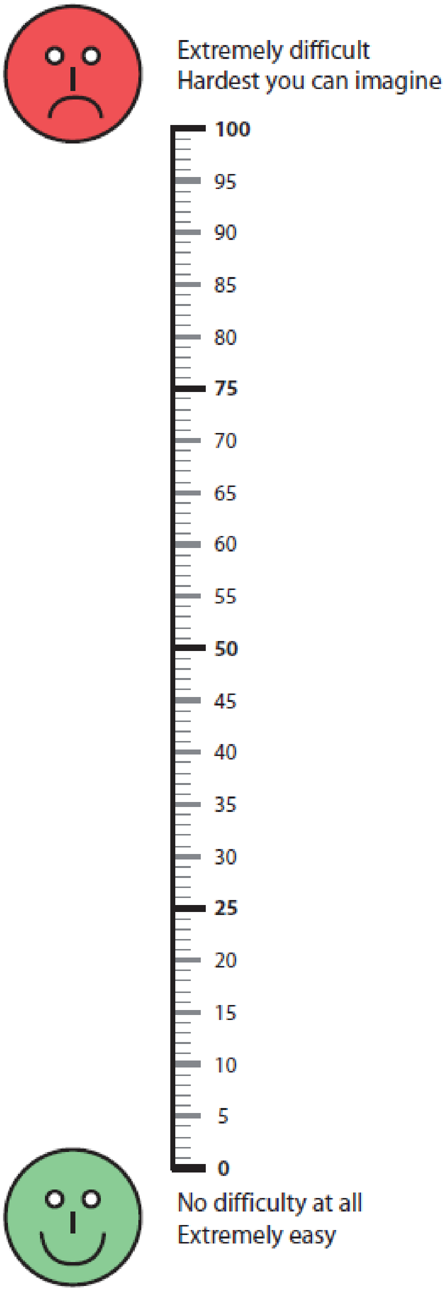
If you have any questions about the study, please contact Umesh Persad at the following address:

Umesh Persad
Cambridge Engineering Design Centre
Department of Engineering
The University of Cambridge
Trumpington Street, Cambridge,
CB2 1PZ, UK

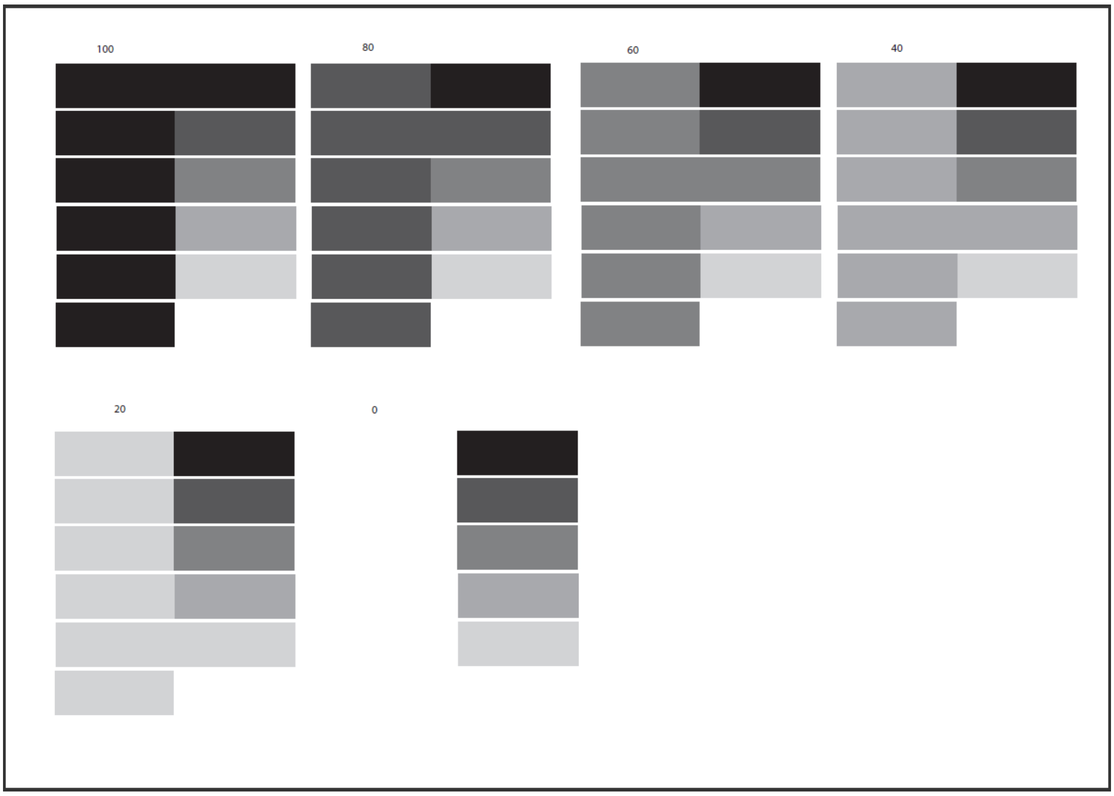
Phone: +44 1223 766958
Mobile: 07984968367
Fax: +44 1223 332662
Email: up209@cam.ac.uk

Appendix 5: Difficulty Scale

Difficulty Scale



Appendix 6: Contrast Cards



Appendix 7: Task Sequences

Clock Radio: Set the time to 4:30PM

Press and hold 'TIME SET' button

While the 'TIME SET' button is held, press the hour button repeatedly (or keep it depressed) to set the hour to 4PM – the PM light should be lit.

While the 'TIME SET" button is held, press the min button repeatedly or keep it depressed) to set the minutes to 30.

Release the 'TIME SET' button when finished.

Mobile Phone: Turn the ringer off via the menu

Press the right top button to select 'Menu'

Scroll through the menu using the centre arrow keys to get to 'Ringer Settings'

Press the right top button to select 'Ringer Settings'

Scroll through the 'Ringer settings' options using the centre arrow keys to get to the 'Off' radio button

Press the right top button to select the 'Off' setting

'All ringer tones switched off' message appears

Press the red 'end call' button to exit out of the menu

Blender: Making a smoothie with a banana and water

Turn power on at the mains

Open top cover of jug

Pour water in

Place banana inside jug

Close top cover of jug

Rotate control to the right to turn the blender on (pulse button could also be used)

Observe blender till blending process is complete

Rotate control anti-clockwise to stop blender from blending

Open top cover

Remove jug from base

Pour into cup

Replace blender jug on base

Replace top cover on jug

Vaccum Cleaner: Cleaning an area of carpet

Extract plug and cord from vacuum cleaner chassis

Plug into mains socket

Turn mains power switch on

Switch on vacuum cleaner by sliding the power switch

Recline handle by grasping the handle and pushing recline button with foot

Optional: Turn on beater bar with a foot push

Place vacuum cleaner over area and push back and forth to clean

When finished, return the handle to the upright position

Slide the power switch to the off position

Unplug the cord from the mains socket

Press retractor button to retract the cord into the vacuum cleaner chassis

Appendix 8: Product Demand Values

Clock Radio Measures	Value	Units
Button text height (time set, hour, minute)	1.5	mm
Button text contrast against chassis	50	%
Button diameter	1.7	mm
Button contrast against chassis	50	%
Sleep button width	12	mm
Sleep button height	7	mm
Digital display numbers text height	14	mm
Digital display numbers text thickness (stroke)	1.5	mm
Digital display numbers text contrast	14	%
PM' text height	2	mm
PM' text contrast against chassis	33	%
Button push force	3.3	N
Viewing distance	40 - 80	cm
Mobile Phone Measures	Value	Units
Button text/numbers height	2	mm
Button text contrast against button	67	%
Button length	5	mm
Button width	11	mm
Button contrast against chassis	60	%
Screen Text (menu and menu items etc.)	2	mm
Screen Text contrast	33	%
OFF' text for taking off phone ringer	4	mm
Weight	0.8	N
Button push force	2.9	N
Blender Measures	Value	Units
Numbers on control height	3	mm
Numbers on control contrast	67	%
"PULSE" text height	3	mm
Rotary control diameter	48	mm
Rotary control contrast against chassis	67	%
Pulse button width	27	mm
Pulse button height	16	mm
Pulse button contrast	67	%
Align Arrows contrast	67	%
Blender cover contrast	100	%
Jug weight	19.3	N
Force to open lid with 1 finger	19.8	N
To pull open from centre	22.9	N
Force to push cover down	176	N
Pulse button push force	9	N
Dial torque clockwise	0.236	Nm
Dial torque anticlockwise	0.177	Nm
Weight of Jug	19.3	N
Setting 1 loudness	82	dBA
Setting 2 loudness	90	dBA
Setting 3 loudness	94	dBA

Vacuum Cleaner Measures	Value	Units
Power button control text height (0, min, max)	3	mm
Power button control text colour	black	
Power button control text contrast	60	%
Beater bar ON/OFF Text height	3	mm
Beater bar ON/OFF Text colour	black	
Beater bar ON/OFF Text contrast	67	%
Slider power button width	27	mm
Slider power button height	14	mm
Slider power button colour	black	
Slider power button contrast	43	%
Cord retract button width	28	mm
Cord retract button height	19	mm
Cord retract button colour	black	
Cord retract button contrast	25	%
Foot push recline button width	44	mm
Foot push recline button height	25	mm
Foot push recline button colour	black	
Foot push recline button contrast against floor	25	%
Loudness Level High	71	dBA
Loudness Level Medium	67	dBA
Loudness Level Low	64	dBA
pull the plug and cord out of cleaner	14.7	N
Slide the on/off switch	13	N
push the handle recline button	9	N
push the beater bar button	6.4	N
push the cleaner forward	15	N
pull the cleaner backward	30	N

Appendix 9: Experimental Study Checklist

Before participant arrives

1. PRODUCTS

□ Ensure the clock radio is **OFF**.

□ Ensure the mobile phone is charged and the ringer is set to **ON**.

□ Ensure there is a *banana* and a *cup of water* set up by the blender.

□ Ensure vacuum cleaner is unplugged in the corner with cord retracted.

□ Ensure there is enough flour to throw for the vacuum task.

□ Products are covered with sheet and vacuum cleaner is behind the curtain.

2. TEST EQUIPMENT

Camera and recorder

□ Ensure the video camera is charged and loaded with a new tape.

□ Ensure I have the audio recorder with enough battery power.

Computer

□ Load up the vision, hearing and cognitive testing software on the computer

□ Enter the current participant ID into CANTAB to run the battery

3. DOCUMENTS

□ Participant Pack: 2 consent forms, boots voucher, questionnaire/data sheet, and task ratings in randomised order.

□ Difficulty scale card is there for the participant

After participant leaves

□ Copy test results from CANTAB onto USB Key (html file) and run a backup.

□ Enter results into spreadsheet

□ Put mobile phone back to charge and make sure the mobile phone ringer is set to ON.

□ Charge force gauge if necessary.

Procedure Guide

Meet participant at the door and take them upstairs.
Introduce myself and give my background.
Thank participant for agreeing to take part.

Go through consent form and emphasise that:

- ➔ We are not testing you, we are testing the product
- ➔ You can take a break at any time you wish
- ➔ You don’t have to answer the questions in the questionnaire if you don’t want to
- ➔ Again we are not testing you
- ➔ Would it be ok if I video you?

Give the participant the Boots voucher
Have participant sign **2 forms.**
Give them one copy and keep one copy.
* Remember to record their email or postal address on the back of my consent form if they want a summary of the study results.

Turn on audio recorder first.
Fill in the questionnaire with all info and mental models while recorder is on.
Take recorder off at the end of the questionnaire.

Take the participant into the usability lab and run the vision, hearing and cognitive tests.

----- BREAK ----- (refreshments)

Run the motor tests following the data sheet.

----- BREAK ----- (refreshments)
While participant is outside, prepare the floor for vacuum task.

Run the participant through the tasks in order in the pack.
Remember to leave audio recorder running throughout.

Take the participant outside and then **debrief** by the table:

➔ Are you still happy for us to use the audio and video recordings?
➔ Explain that the purpose of the study was to look at matching up the demands of the products to your capabilities to see how we can improve the design to better match up to you.
➔Please don't tell any one else (especially other if you know other volunteers) what you did during the session as that will bias the results.
➔ Arrange for transportation costs to be refunded if needed (with cash or the form).
➔ Are you happy with everything? Good.

Transport participant back if necessary.

Appendix 10: Experiment Data Collection Sheets

Participant Questionnaire

Participant ID:	
Date of session:	
Session start time:	
Session end time:	
Session duration:	

REMINDER:
Turn on audio recorder at the start of this session

Demographic Information

Age: ……….

Gender: ☐ Male ☐ Female

Highest Education Level:

…………………………………………………………………………………………………..

Current Occupation:

…………………………………………………………………………………………………..

Health

Do you have any long term health problems or complaints that affect your everyday activities? If so, please list up to 4 current health/medical conditions:

1……………………………………………………………………………………………………

2……………………………………………………………………………………………………

3……………………………………………………………………………………………………

4……………………………………………………………………………………………………

General capability

1. Do you currently have any problems with your vision that affects you in your everyday activities?
□ Yes □ No

If yes, please describe the nature of these problems:

……………………………………………………………………………………………………

……………………………………………………………………………………………………

Do you currently use any aids to help with seeing? (e.g. magnifiers)

……………………………………………………………………………………………………

2. Do you have any problems with your hearing that affects you in your everyday activities?
□ Yes □ No

If yes, please describe the nature of these problems:

……………………………………………………………………………………………………

……………………………………………………………………………………………………

Do you currently use any aids to help with hearing? (e.g. hearing aids)

……………………………………………………………………………………………………

3. Do you have problems with your memory that affects you in your everyday activities?
□ Yes □ No

If yes, please describe the nature of these problems:

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

4. Do you currently have any problems with the use of your hands and arms that affects you in your everyday activities?
□ Yes □ No

If yes, please describe the nature of these problems:

……………………………………………………………………………………………………

……

5. Do you currently have any problems walking and moving about normally?
□ Yes □ No

If yes, please describe the nature of these problems:

……

……

Do you currently use any aids to help with moving about? (e.g. walker/wheelchair/ cane)

……

General Product Experience

Clock Radio Experience

Do you currently own a clock radio (or alarm clock)? □ Yes □ No

How often do you personally use a clock radio (*never, more than once a day, more than once a week, more than once a month, more than once a year*)?

……………………………………………………………………………………………………

Could you describe the kind of clock radio or alarm clock that you currently use (or most recently used)? [digital or analogue, brand]

……………………………………………………………………………………………………

How would you rate your experience with clock radios on a scale from 0 to 10, where 0 represents no experience at all and 10 represents a very high level of experience:

……………………………………………………………………………………………………

Mobile Phone Experience

Do you currently own a mobile phone? □ Yes □ No

How often do you personally use a mobile phone (*never, more than once a day, more than once a week, more than once a month, more than once a year*)?

……………………………………………………………………………………………………

Could you describe the kind of mobile phone that you currently use (or most recently used)? [flip or candy bar, brand]

……………………………………………………………………………………………………

How would you rate your experience with mobile phones on a scale from 0 to 10, where 0 represents no experience at all and 10 represents a very high level of experience:

……………………………………………………………………………………………………

Blender Experience

Do you currently own a blender? □ Yes □ No

How often do you personally use a blender (*never, more than once a day, more than once a week, more than once a month, more than once a year*)?

……………………………………………………………………………………………………

Could you describe the kind of blender that you currently use (or most recently used)? [brand, type]

……………………………………………………………………………………………………

How would you rate your experience with blenders on a scale from 0 to 10, where 0 represents no experience at all and 10 represents a very high level of experience:

……………………………………………………………………………………………………

Vacuum Cleaner Experience

Do you currently own an upright vacuum cleaner? □ Yes □ No

……………………………………………………………………………………………………

How often do you personally use a vacuum cleaner (*never, more than once a day, more than once a week, more than once a month, more than once a year*)?

……………………………………………………………………………………………………

Could you describe the kind of vacuum cleaner that you currently use (or most recently used)? [brand, upright or cylinder]

……………………………………………………………………………………………………

How would you rate your experience with vacuum cleaners on a scale from 0 to 10, where 0 represents no experience at all and 10 represents a very high level of experience:

……………………………………………………………………………………………………

Specific Task and Product Experience

"I am now going to ask you to tell me how you would go about using some products. I would like you to tell me, in as much detail as you could, the steps you would go through and which features of the product you will use. Ok?"

1. Please tell me, in as much detail as you could, the steps you would go through and the product features you will use when ***setting the time on a digital clock-radio.***

2. Please tell me, in as much detail as you could, the steps you would go through and the product features you will use when ***taking the ringer off a mobile phone.***

3. Please tell me, in as much detail as you could, the steps you would go through and the product features you will use when ***using a blender to make a drink or smoothie (or just blending/mixing).***

4. Please tell me, in as much detail as you could, the steps you would go through and the product features you will use when ***using an upright vacuum cleaner to vacuum clean a piece of carpet.***

Probes:
Could you explain in some more detail?
Tell me more about the individual steps and which specific features you would use.

REMINDER:
Turn off audio recorder at the end of this session

Part 2 Capability Data

Vision and Hearing

1. Distance Acuity

Seat participant 3m from the screen and turn down the lights (dimmer)
***Click** on LogMAR test, then **click** on Random*
Ask participant to read from the top and count the letters correctly identified (VAR)

Distance Acuity ……….. logMAR ……….. VAR

Could you pick a line above that represents when the size of the text starts to become difficult to read?

Comfort level ……….. logMAR ……….. VAR

2. Contrast Sensitivity

Seat participant 1m from screen (measure 1m distance with string)
Ask participant not to move or lean forward
Click on CS test and start from 1.8 logMAR
Tell participant to read from top of screen, and rows will appear below. Continue reading till you cannot see the letters.

LogMAR at 1m	Log Contrast Sensitivity	% Contrast
1.8		
1.6		
1.3		
1.0		
0.7		
0.4		

3. Near/reading Acuity

Seat participant 40 from the screen, measure with string

Reading Acuity ……….. logMAR ……….. VAR

Comfort level ……….. logMAR ……….. VAR

4. Hearing level

Seat 1m from the screen, measure with string

Can you hear this sound? (Low, 2^{nd} bar up, 57dB) □ Yes □ No
Can you hear this sound? (Medium, 4^{th} bar up, 67dB) □ Yes □ No
Can you hear this sound? (High, top bar 73dB) □ Yes □ No

REMEMBER TO SET THE VOLUME SLIDER TO MID LEVEL FOR COG TESTS

Cognition

Say to participant:

"I am going to read out a series of numbers. I would like you to repeat them back to me in the same order. I will increase the length of the series as we go along."

Digit span test

Digit span is 1	9 4 3
Digit span is 2	1 3 4 2 1 5
Digit span is 3	8 7 8 3 8 8 3 1 9
Digit span is 4	9 7 5 4 3 8 2 5 6 5 1 4
Digit span is 5	9 4 3 1 8 6 8 2 5 9 3 8 1 4 7
Digit span is 6	9 1 3 8 2 5 6 4 8 3 7 1 5 9 6 3 8 2
Digit span is 7	7 9 5 8 4 2 3 5 3 1 6 8 4 2 7 9 1 8 5 4 6
Digit span is 8	8 6 9 5 1 3 7 2 5 1 7 3 9 8 2 6 5 1 3 9 8 2 4 7
Digit span is 9	7 1 9 3 8 4 2 6 1 1 6 3 8 7 4 9 5 2 6 2 5 9 4 3 8 2 6
Digit span is 10	9 1 5 2 4 3 8 1 6 2 7 1 5 4 8 5 6 1 9 3 1 5 2 8 4 6 7 3 1 8

Digit span is 1	4 2 5
Digit span is 2	2 1 2 7 7 4
Digit span is 3	9 5 1 6 1 5 9 2 3
Digit span is 4	4 8 5 7 1 9 3 6 1 7 3 2
Digit span is 5	9 1 6 3 8 8 9 2 4 3 4 7 2 1 3
Digit span is 6	8 2 7 3 1 6 6 3 4 5 8 7 5 3 2 4 7 1
Digit span is 7	3 9 4 6 2 7 8 2 3 7 9 1 5 6 4 7 1 9 6 3 2
Digit span is 8	8 2 5 7 9 6 3 1 8 2 3 1 9 4 6 5 1 2 6 3 4 8 9 5
Digit span is 9	7 5 4 1 6 9 2 3 8 9 6 1 4 7 3 8 5 2 8 6 5 3 1 7 4 9 2
Digit span is 10	9 8 2 1 3 6 4 1 9 7 2 7 5 8 4 9 6 9 5 1 8 5 4 8 6 5 8 6 1 6

CANTAB Tests

Administer CANTAB tests. Say to participant: "Now we are going to simple tasks with a computer. I will give you instructions as we go along."

Enter participant number into CANTAB and start the battery:

2. Motor Screen
3. Spatial Span
4. Reaction time
5. Graded naming test

- - - BREAK - - -

Motor Tests

Say to Participants:
Now I am going take some physical measures like the forces you can exert with your hands. For measures of comfort, I want you to exert force till it starts to become difficult, then hold for two seconds, and then release. For maximum forces, gradually exert force as much as you can, hold for two seconds, then release. Ok?

Preferred Hand

What is your preferred hand? ☐ Right ☐ Left ☐ Any

1. Grip Strength, (Jamar Hand Gauge)

Graduated in 2kg force units. Rest gauge on participant leg with me holding it, elbow at 90 degrees..

Right Hand Comfortable ……….. Kg Force

Left Hand Comfortable ……….. Kg Force

Right Hand Max (1st) ……….. Kg Force

Left Hand Max (1st) ……….. Kg Force

2. Thumb/index finger, lateral pinch grip (Jamar Pinch Gauge)

Graduated in 0.5kg force units. Rest gauge flat on table and use lateral thumb pinch.

Right Hand Comfortable ……….. Kg Force

Left Hand Comfortable ……….. Kg Force

Right Hand Max (1st) ……….. Kg Force

Left Hand Max (1st) ……….. Kg Force

3. Rotational torque, lateral pinch grip, (Torque screwdriver)

Right Hand Clockwise Comfortable ……….. Nm

Right Hand Anti-Clockwise Comfortable ……….. Nm

Left Hand Clockwise Comfortable ……….. Nm

Left Hand Anti-Clockwise Comfortable ……….. Nm

Right Hand Clockwise Max ……….. Nm

Right Hand Anti-Clockwise Max ……….. Nm

Left Hand Clockwise Max ……….. Nm

Left Hand Anti-Clockwise Max ……….. Nm

4. Index Finger and Thumb Push

Remove torque sensor and attach the plate to the end of force gauge. Participant seated in front of gauge.

Right Hand index finger push comfortable ……….. N **R**Thumb ……….. N

Left Hand index finger push comfortable ……….. N LThumb ……….. N

Right Hand index finger push force Max ……….. N **R**Thumb ……….. N

Left Hand index finger push force Max ……….. N LThumb ……….. N

5. Pulling with power grasp

Attach the hook to the end of force gauge. Participant standing with bar to his/her side

Right Hand pull force comfortable ……….. N

Left Hand pull force comfortable ……….. N

Right Hand pull force Max ……….. N

Left Hand pull force Max ……….. N

6. Pushing with power grasp

Turn participant around with bar to the side and pull forward simulating a push

Right Hand push force comfortable ……….. N

Left Hand push force comfortable ……….. N

Right Hand push force Max ……….. N

Left Hand push force Max ……….. N

7. Vertical lift force

Rotate gauge on table, and sit on table

Right Hand lift force comfortable……….. N

Left Hand lift force comfortable ……….. N

Right Hand lift force Max ……….. N

Left Hand lift force Max ……….. N

8. Two Finger pull-up force

Remove baton from the hook and ask participant to use index and middle fingers of preferred hand to pull up.

2 finger pull-up force comfortable ……. N

2 finger pull-up force Max ……. N

9. Balance time

Balance time on preferred leg ……….. Sec

10. Walking speed

Record the time to walk and touch the wall and back (in the corridor).

Time to walk 7.8m comfortably ……….. Sec

Time to walk 7.8m hurried ……….. Sec

- - - BREAK - - -

Product Tasks

Participant Number: ....

Reminders

1. Uncover products

2. Sprinkle flour on floor

3. Put vacuum cleaner in corner

4. Ensure banana and water is available

4. Make sure mobile phone ringer is off and it is charged

REMINDER:
Turn on audio recorder at the start of this session

1 Clock Radio Task

Start videotape and instruct to perform the task:
Could you please set the time on the clock to 4:30 PM? Please tell me when you are finished with the task, ok?

Difficulty Ratings:
Give participant the scale and say: "I would like you to use this scale to rate your difficulty with various actions"

Table 1 Difficulty scores for the clock radio

Action	Scores
Vision	
How easy/difficult was it to **see** the buttons that you used?	
How easy/difficult was it to **see/read** the button text labels?	
How easy/difficult was it to **see/read** the numbers on the digital display?	
How easy/difficult was it to **see/read** the text 'PM' on the display?	
Physical actions	
How easy/difficult was it to **push** the individual buttons with your fingers?	
How physically demanding was it to perform this task overall?	
Mental actions	
How difficult was it to work out what to do when you were starting the task of setting the time?	
While setting the time, how difficult was it to work out what action to do next?	
How **mentally demanding/difficult** did you find this task overall?	

Difficulty experienced

Could you explain which parts of this task you found to be the most difficult and how the product caused your difficulty?
(Probe: Any other actions/part of the task gave you trouble?)

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......
……………………………………………………………………………………………......
……………………………………………………………………………………………......

Frustration

Did you become frustrated at any point during the task? □ Yes □ No

At which point were you the most frustrated?

……………………………………………………………………………………………......
……………………………………………………………………………………………......
……………………………………………………………………………………………......
……………………………………………………………………………………………......
……………………………………………………………………………………………......

Please rate your level of frustration on a scale from 0 to 100:

……………………………………………………………………………………………......

□ Task not attempted, □ Successful completion, □ Thought goal achieved, but failed □ Gave up during task

2 Mobile Phone Task

Start videotape and instruct to perform the task:
This phone has a setting to take the ringer off, so that it will not make a sound when someone is calling. Could you take the ringer off?
Please tell me when you are finished with the task, ok?

Difficulty Ratings:
Give participant the scale and say: "I would like you to use this scale to rate your difficulty with various actions"

Table 2 Difficulty scores for the mobile phone

Action	Scores
Vision	
How easy/difficult was it to **see** the buttons?	
How easy/difficult was it to **see/read** the text and numbers on the buttons?	
How easy/difficult was it to **see/read** the text on the display screen?	
Physical actions	
How easy/difficult was it to **grasp and hold** the phone?	
How easy/difficult was it to **press** the buttons on the phone?	
How physically demanding was it to perform this task overall?	
Mental actions	
How difficult was it to work out what to do when you were starting the task of turning off the ringer on the phone?	
While performing the task of taking the ringer off, how difficult was it to work out what action to do next?	
How **mentally demanding/difficult** did you find this task overall?	

Difficulty experienced

Could you explain which parts of this task you found to be the most difficult and how the product caused your difficulty?
(Probe: Any other actions/part of the task gave you trouble?)

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

Frustration

Did you become frustrated at any point during the task? □ Yes □ No

At which point were you the most frustrated?

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

Please rate your level of frustration on a scale from 0 to 100:

…………………………………………………………………………………………………......

□ Task not attempted, □ Successful completion, □ Thought goal achieved, but failed □ Gave up during task

3 Blender Task

Start videotape and instruct to perform the task:
I have a banana and some water on the table. Could you use the blender to blend them make a drink? When you are finished, pour the drink into the cup. Please tell me when you are finished with the task, ok?

Difficulty Ratings:
Give participant the scale and say: "I would like you to use this scale to rate your difficulty with various actions"

Table 3 Difficulty scores for the blender

Action	Scores
Vision	
How easy/difficult was it to **see** the blender control dial?	
How easy/difficult was it to **see/read** the numbers around the control dial?	
Hearing	
How easy/difficult was it to **hear** the blender working when it was on?	
Physical actions	
How easy/difficult was it to **pull open** the cover of the blender jug?	
How easy/difficult was it to **close** the cover of the blender jug?	
How easy/difficult was it to **turn** the control dial?	
How easy/difficult was it to **push** the *PULSE* button?	
How easy/difficult was it to **lift** the blender jug?	
How easy/difficult was it to **pour** from the blender jug into the cup?	
How **physically demanding** was it to perform this task overall?	
Mental actions	
How difficult was it to work out what to do when you were initially starting the task of blending the drink?	
While using the blender, how difficult was it to work out what action to do next?	
How **mentally demanding/difficult** did you find this task overall?	

Difficulty experienced

Could you explain which parts of this task you found to be the most difficult and how the product caused your difficulty?
(Probe: Any other actions/part of the task gave you trouble?)

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

…………………………………………………………………………………………………......

……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......

Frustration

Did you become frustrated at any point during the task? □ Yes □ No

At which point were you the most frustrated?

……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......
……………………………………………………………………………………………………......

Please rate your level of frustration on a scale from 0 to 100:

……………………………………………………………………………………………………......

□ Task not attempted, □ Successful completion, □ Thought goal achieved, but failed □ Gave up during task

4 Vacuum Cleaner Task

Start videotape and instruct to perform the task:
Could you use the vacuum cleaner to clean this piece of carpet till it is clean? When you are done, return the vacuum cleaner to the corner as you found it.
Please tell me when you are finished with the task, ok?

Difficulty Ratings:
Give participant the scale and say: "I would like you to use this scale to rate your difficulty with various actions"

Table 4 Difficulty scores for the mobile phone

Action	Scores
Vision	
How easy/difficult was it to **see** the on/off switch on the vacuum cleaner?	
How easy/difficult was it to **see/read** the text on the *on/off* switch?	
How easy/difficult was it to **see** the handle recline button?	
How easy/difficult was it to **see** the beater bar button?	
How easy/difficult was it to **see/read** the text label for the beater bar button?	
How easy/difficult was it to **see** the cord pull/retractor button?	
Hearing	
How easy/difficult was it to **hear** the vacuum cleaner working when it was on?	
Physical actions	
How easy/difficult was it to **pull** the plug and cord out?	
How easy/difficult was it to **slide** the on/off switch?	
How easy/difficult was it to **push** the handle recline button?	
How easy/difficult was it to **push** the beater bar button?	
How easy/difficult was it to **push** the cleaner forward?	
How easy/difficult was it to **pull** the cleaner backward?	
How **physically demanding** did you find this task overall?	
Mental actions	
How difficult was it to figure out what to do when you were starting the task of vacuuming the carpet?	
While using the vacuum cleaner, how difficult was it to work out what action to do next?	
How mentally demanding/difficult did you find this task overall?	

Difficulty experienced

Could you explain which parts of this task you found to be the most difficult and how the product caused your difficulty?
(Probe: Any other actions/part of the task gave you trouble?)

………………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

Frustration

Did you become frustrated at any point during the task? □ Yes □ No

At which point were you the most frustrated?

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

……………………………………………………………………………………………......

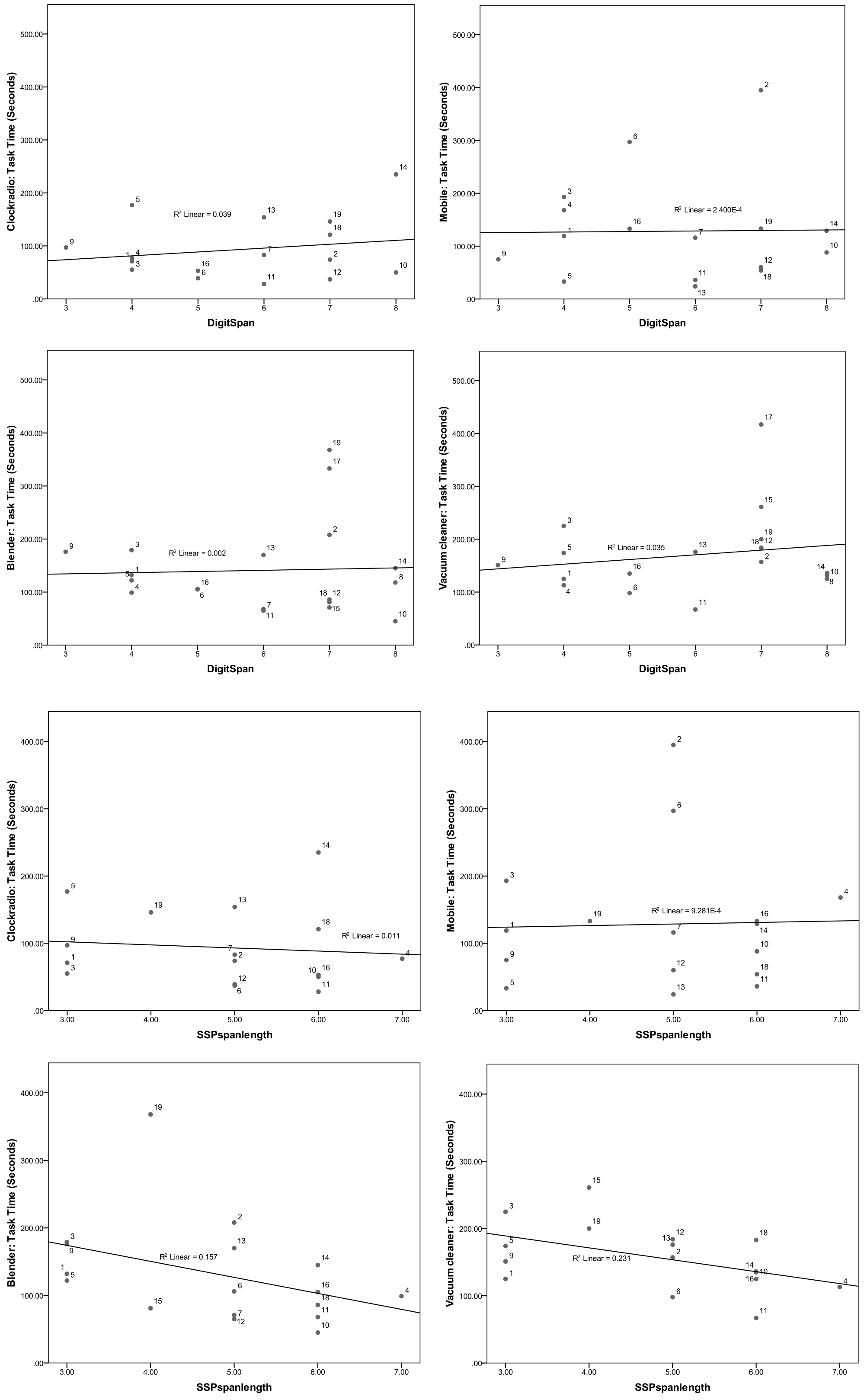
……………………………………………………………………………………………......

Please rate your level of frustration on a scale from 0 to 100:

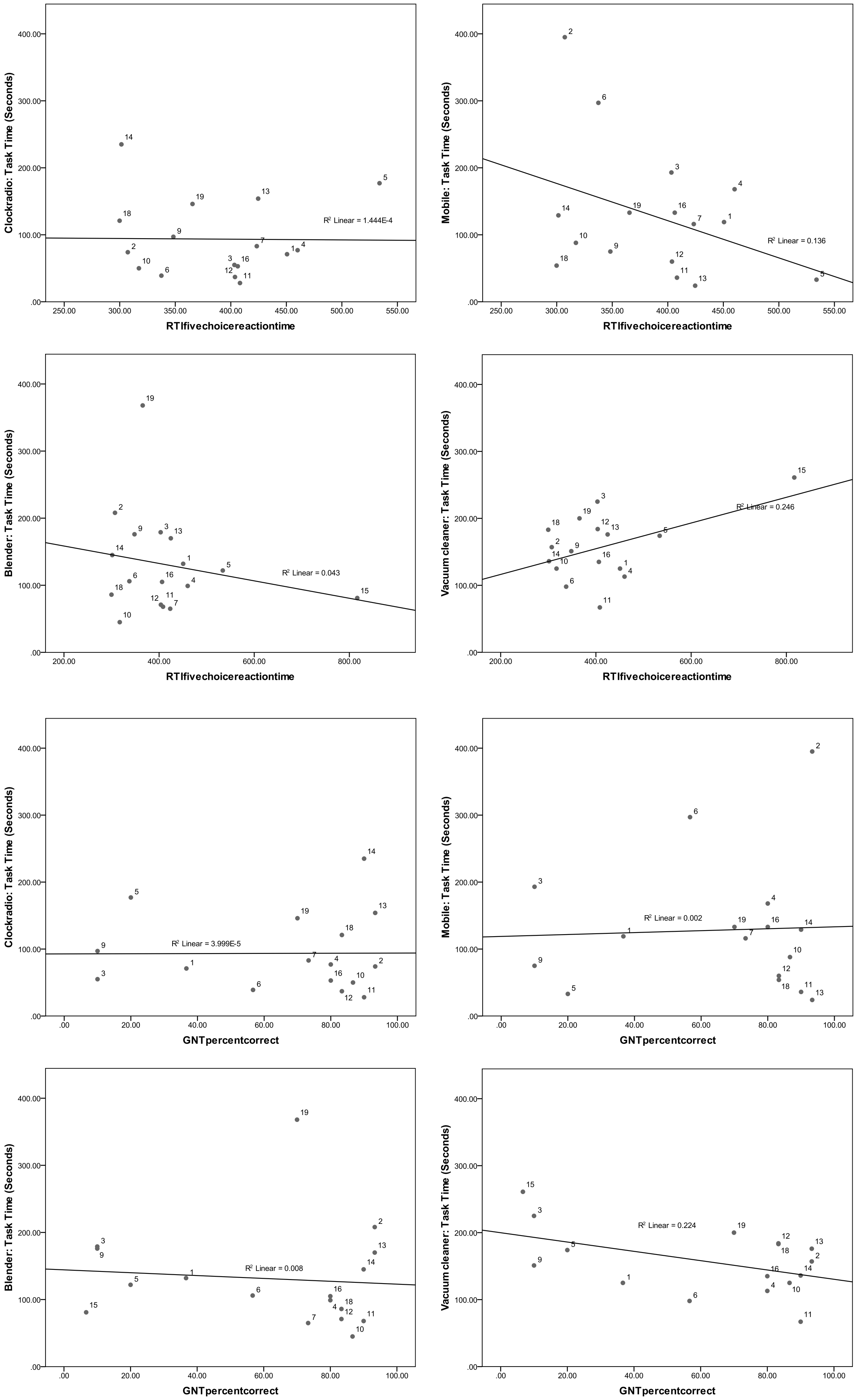
……………………………………………………………………………………………......

□ Task not attempted, □ Successful completion, □ Thought goal achieved, but failed □ Gave up during task

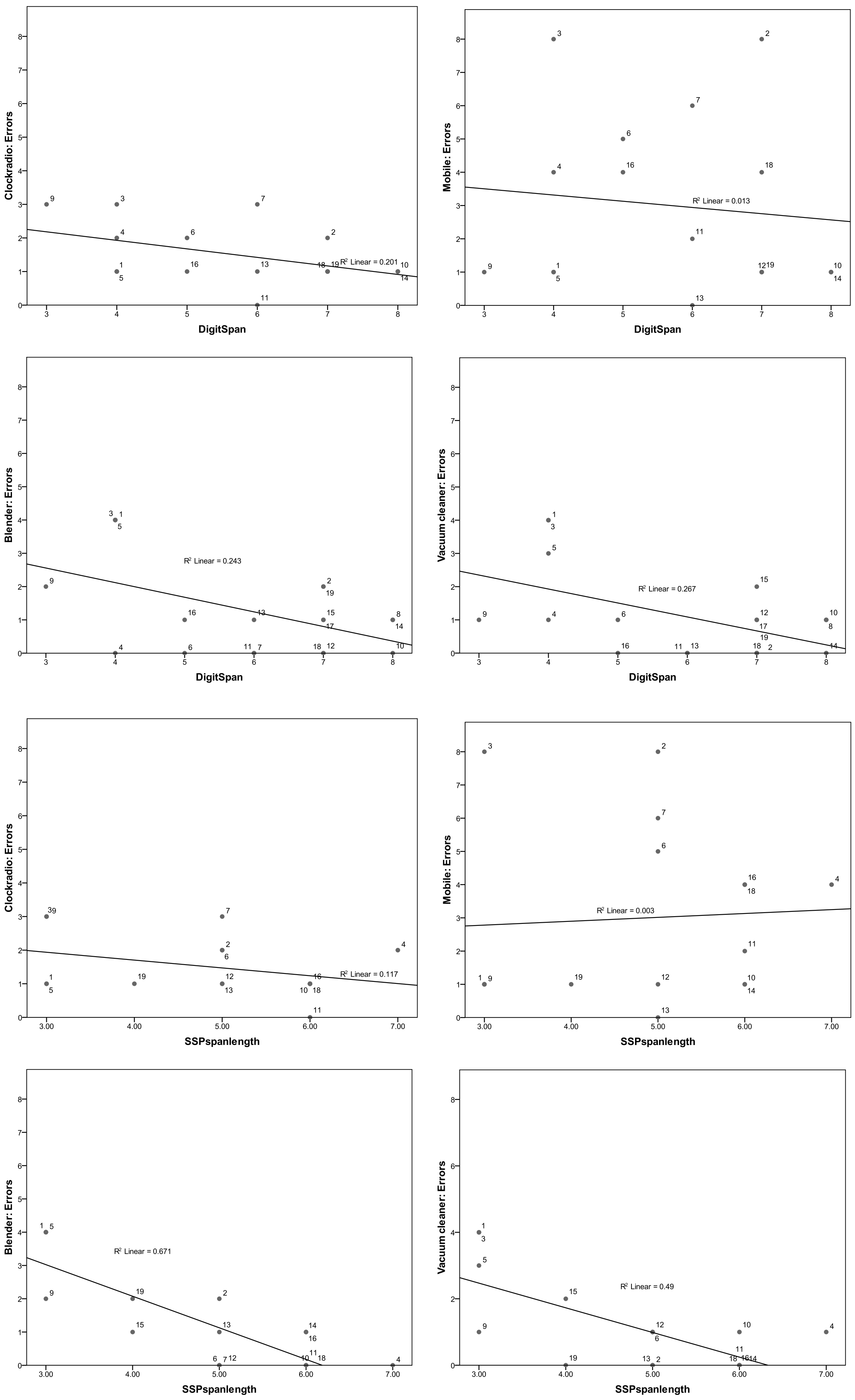
Appendix 11: Scatter Plots



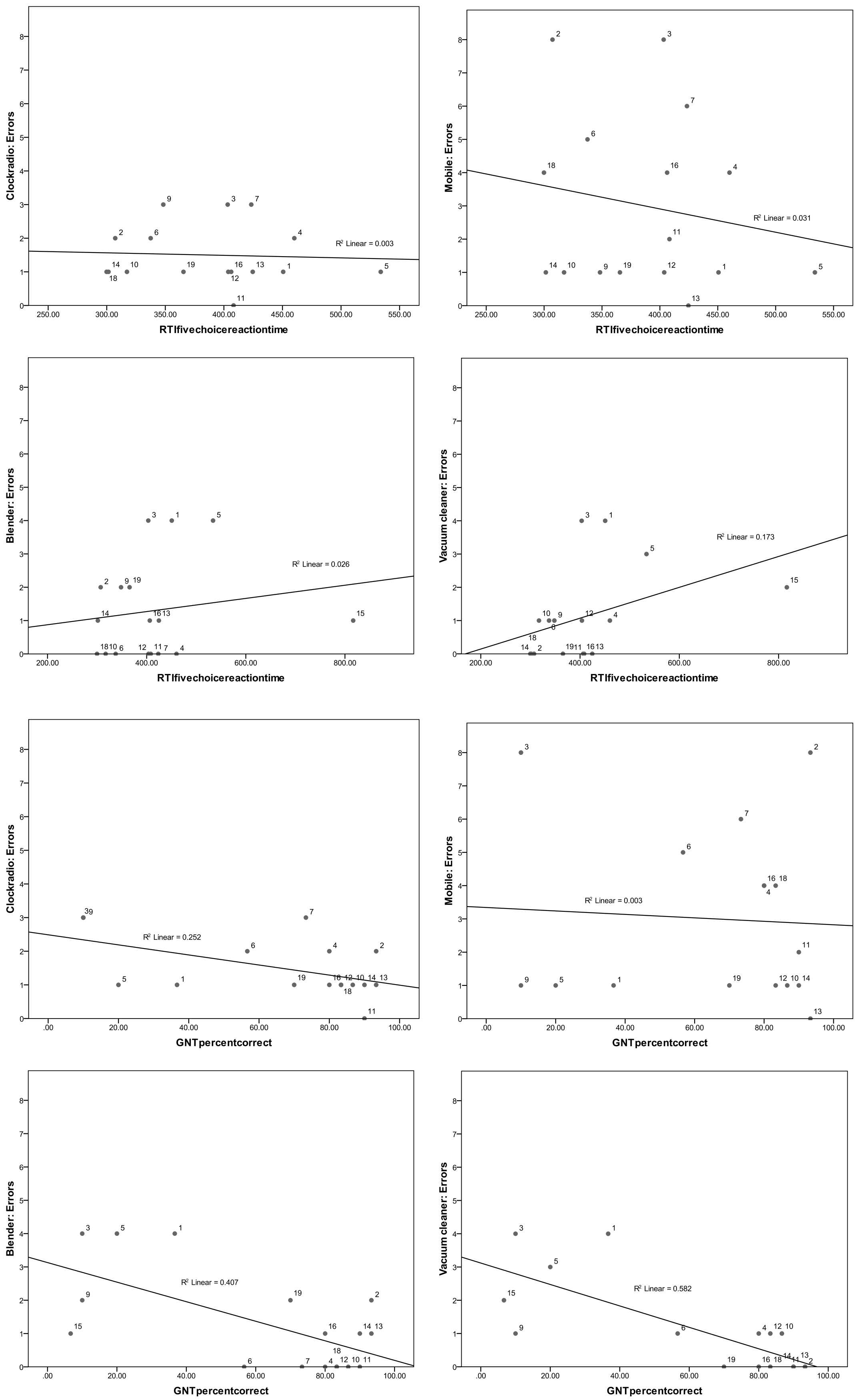
Scatter plots of task time versus digit span and span length



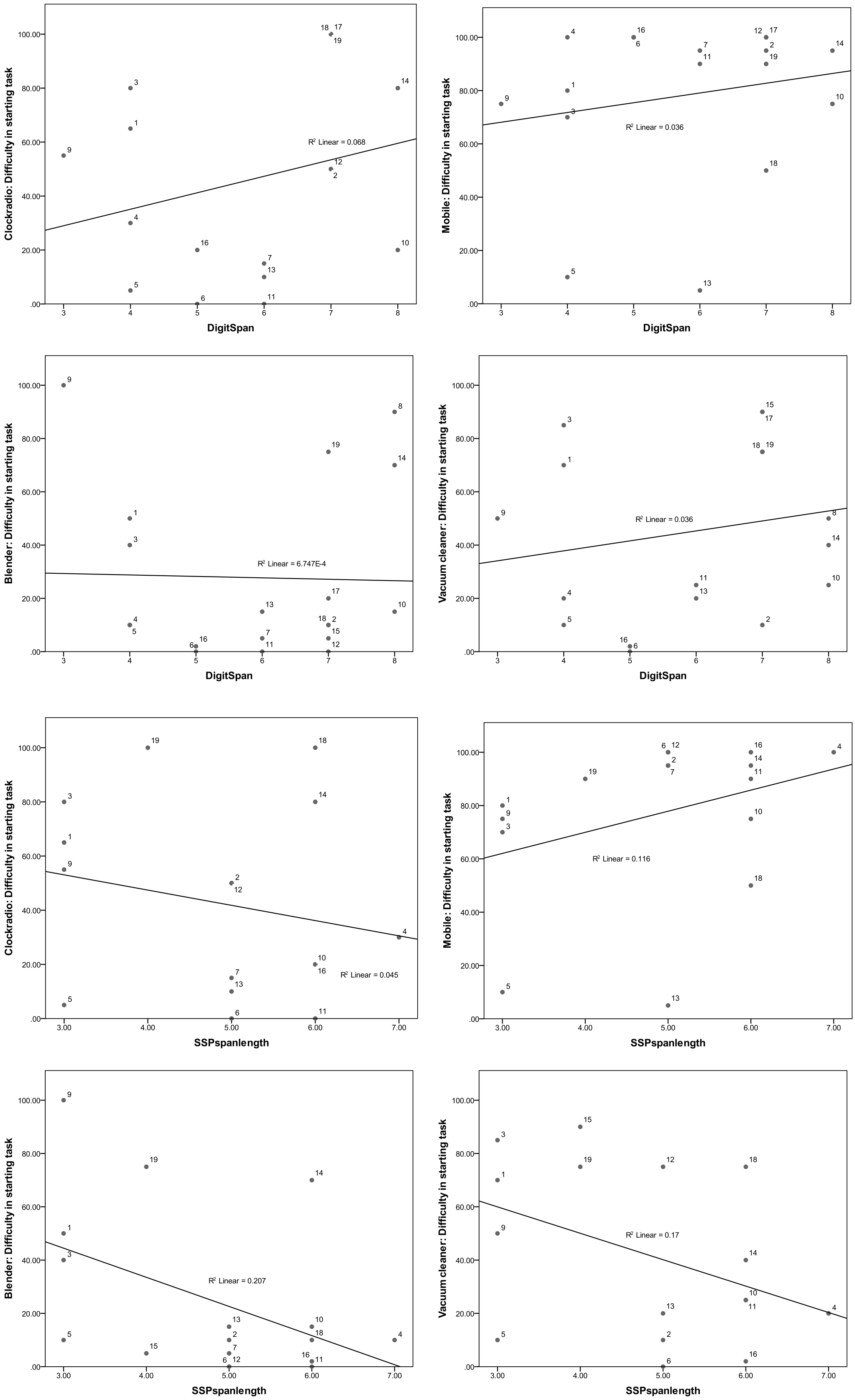
Scatter plots of task time versus reaction time and GNTpercentcorrect



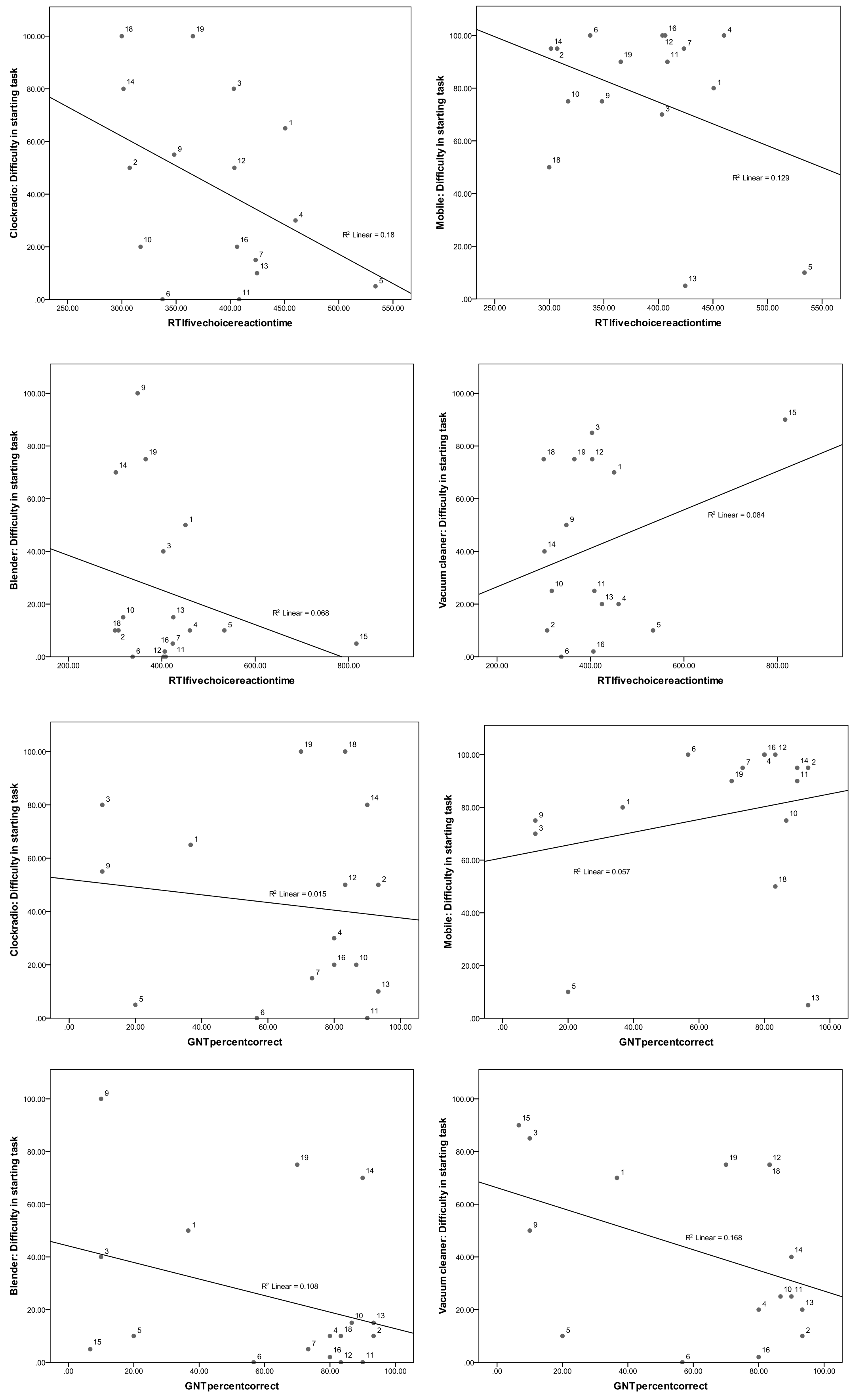
Scatter plots of number of errors versus digit span and span length



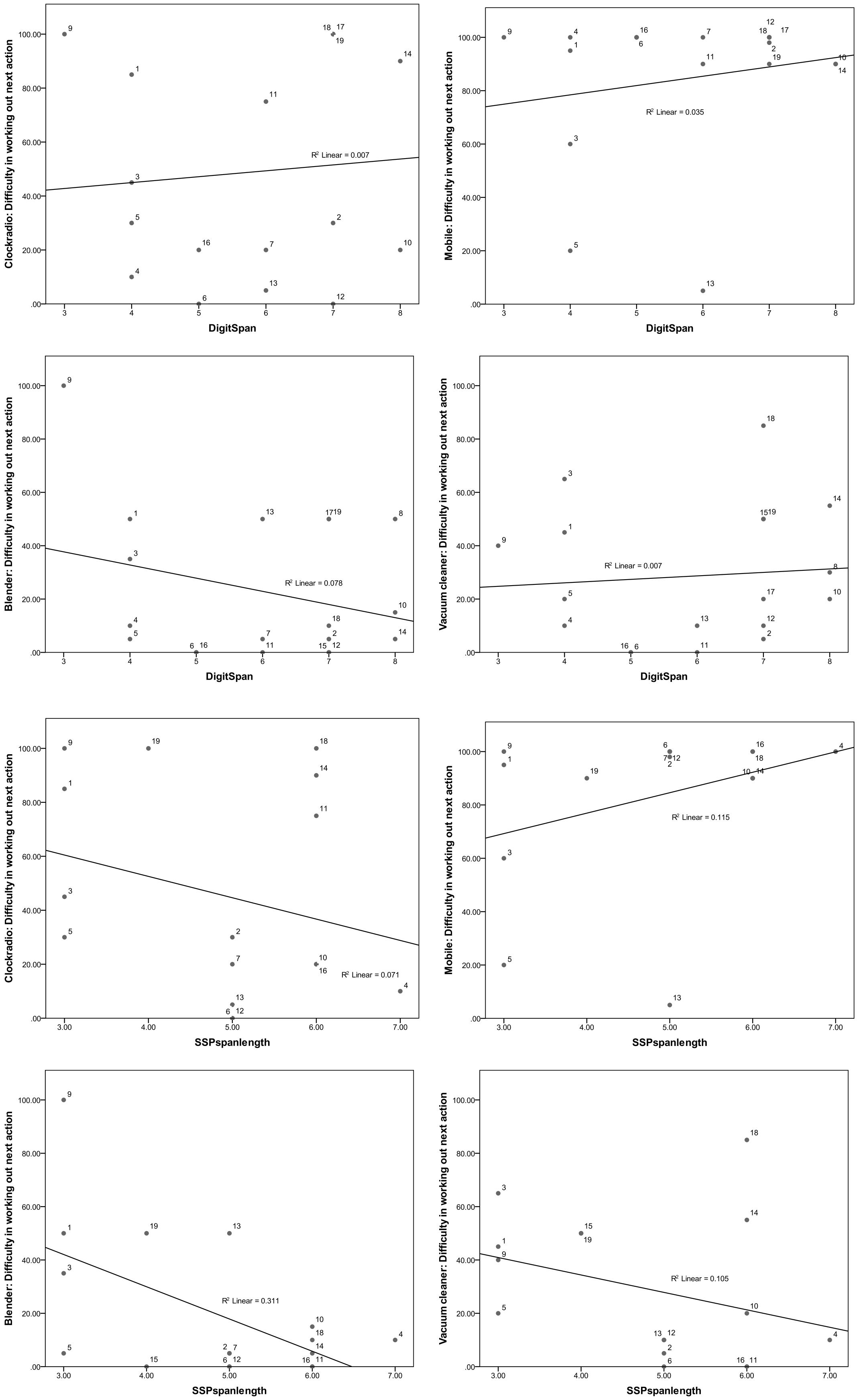
Scatter plots of number of errors versus reaction time and GNTpercentcorrect



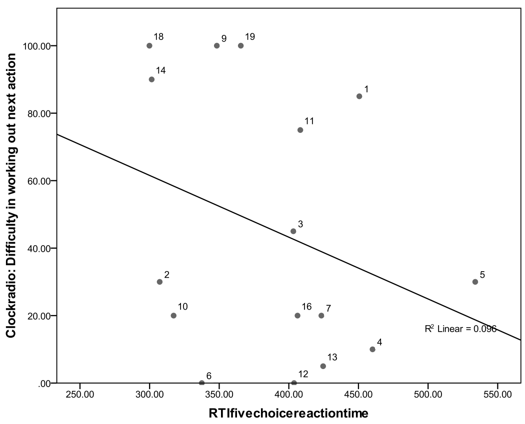
Scatter plots of difficulty starting task versus digit span and span length

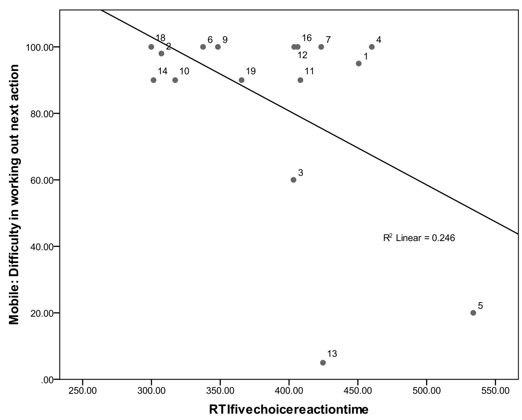


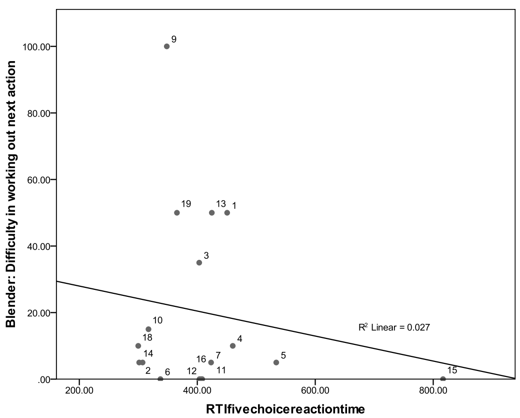
Scatter plots of difficulty starting task versus reaction time and GNTpercentcorrect

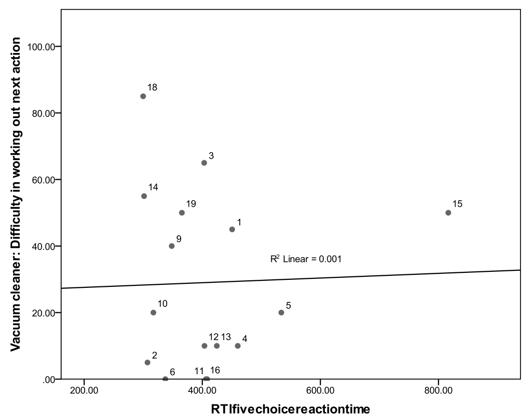


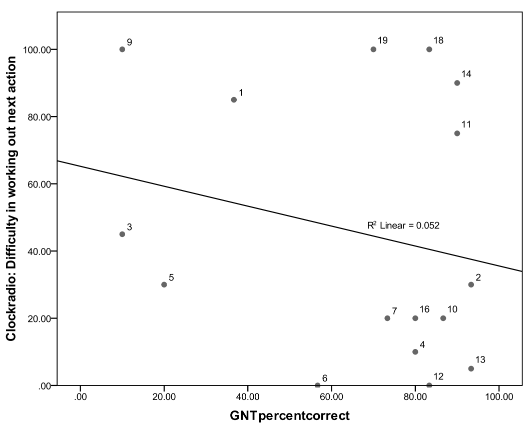
Scatter plots of difficulty in working out next action versus digit span and span length

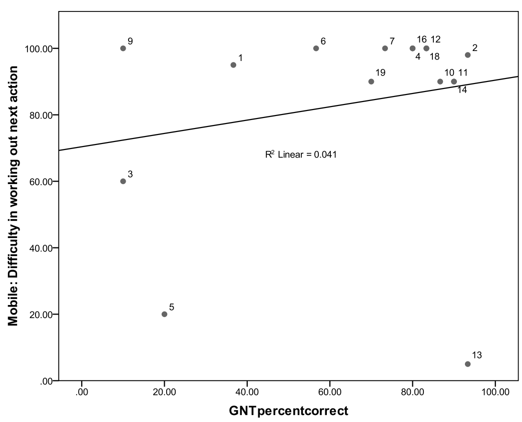

Clockradio: Difficulty in working out next action
RTIfivechoicereactiontime
R² Linear = 0.096

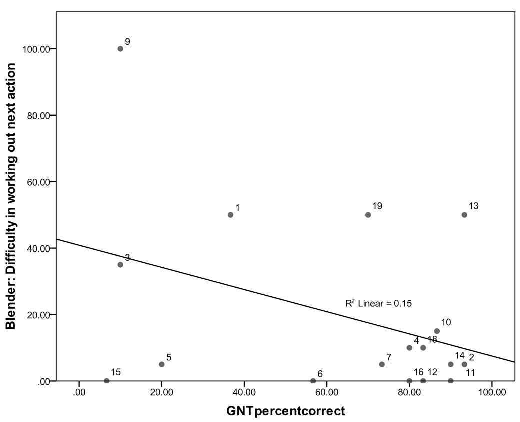

Mobile: Difficulty in working out next action
RTIfivechoicereactiontime
R² Linear = 0.246

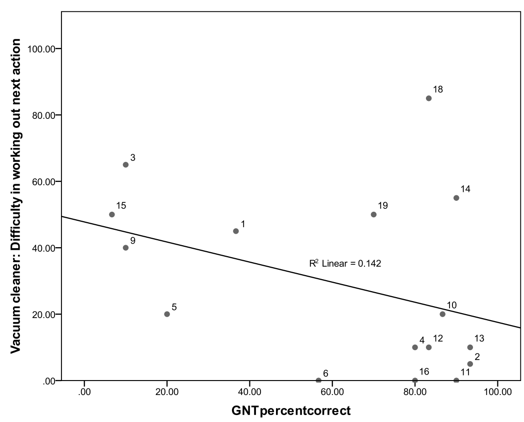

Blender: Difficulty in working out next action
RTIfivechoicereactiontime
R² Linear = 0.027


Vacuum cleaner: Difficulty in working out next action
RTIfivechoicereactiontime
R² Linear = 0.001

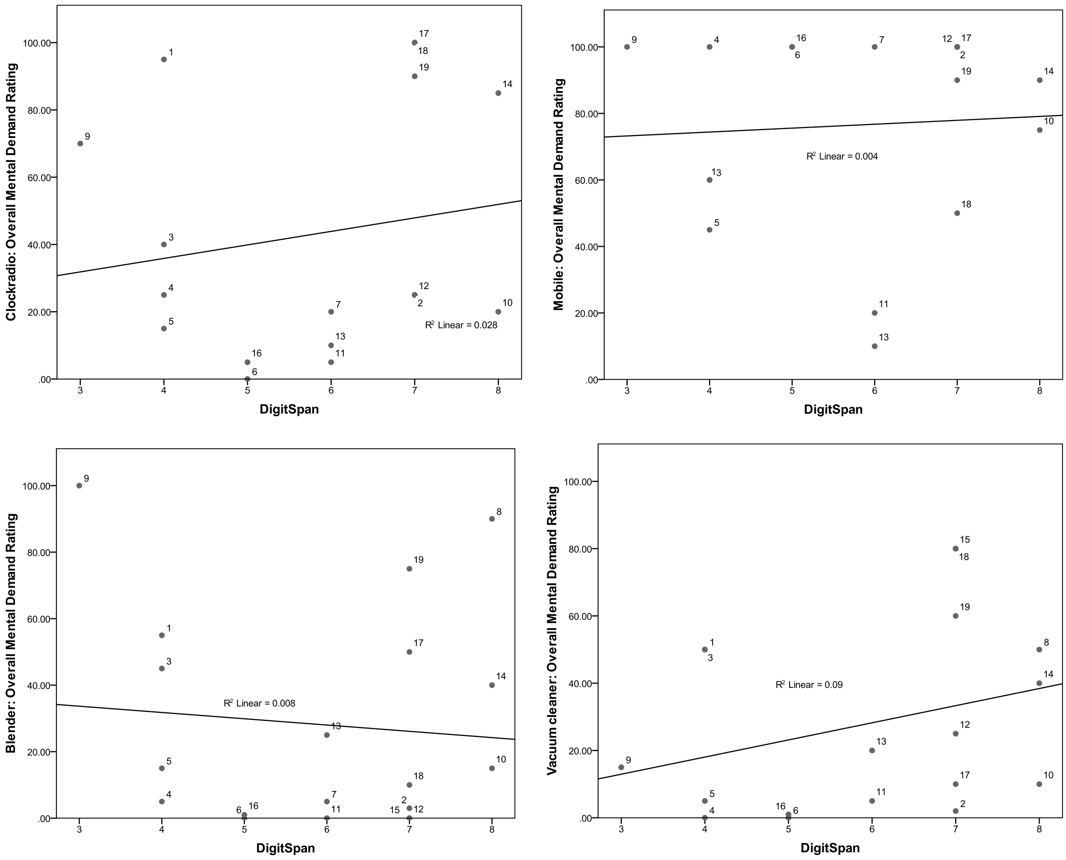

Clockradio: Difficulty in working out next action
GNTpercentcorrect
R² Linear = 0.052

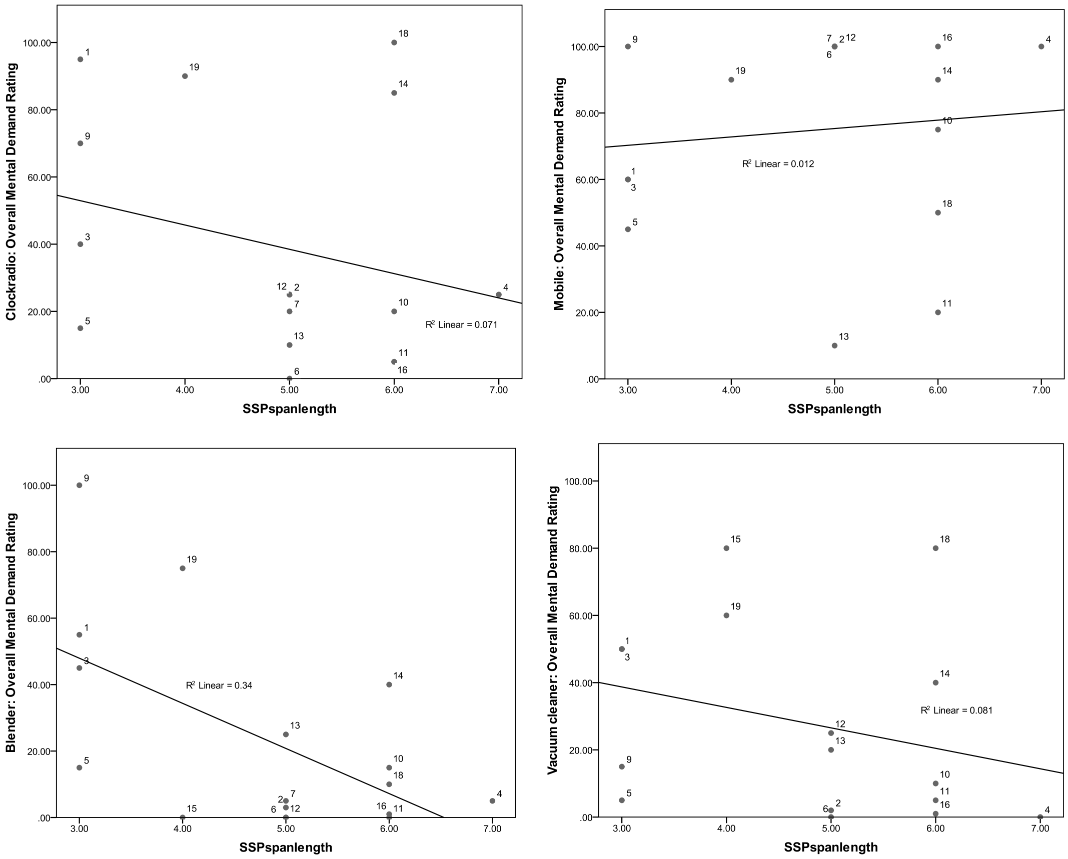

Mobile: Difficulty in working out next action
GNTpercentcorrect
R² Linear = 0.041


Blender: Difficulty in working out next action
GNTpercentcorrect
R² Linear = 0.15

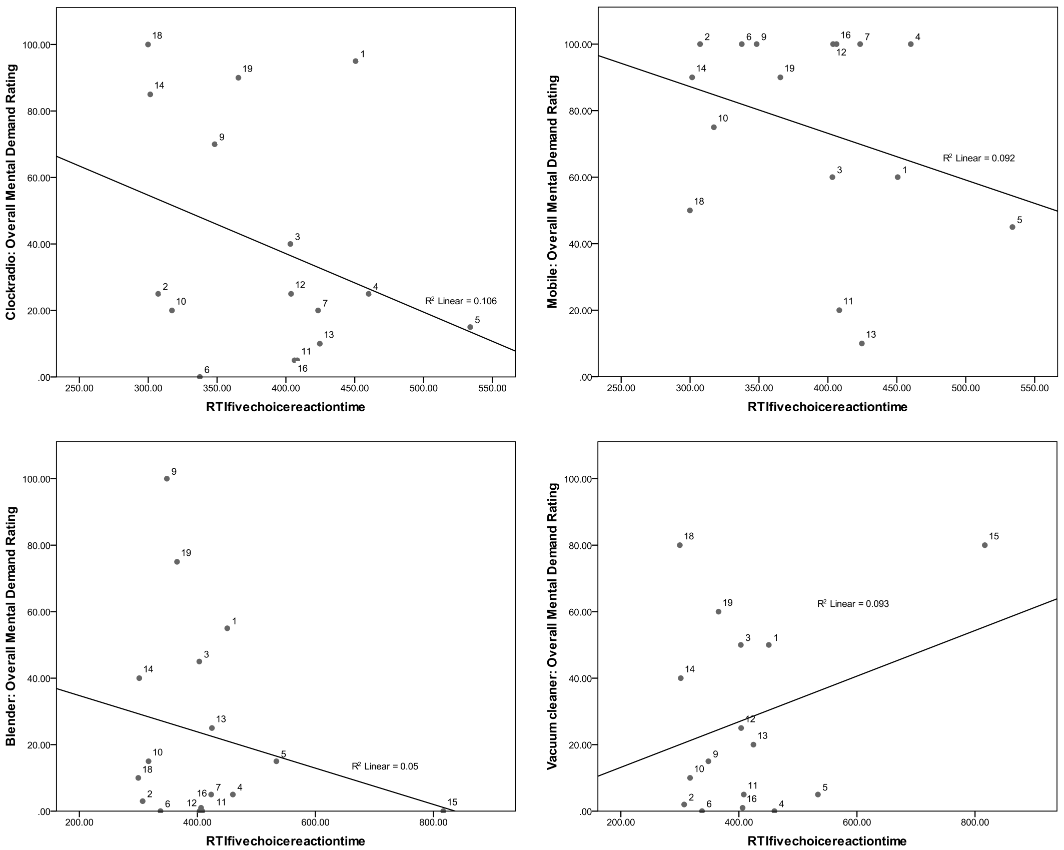

Vacuum cleaner: Difficulty in working out next action
GNTpercentcorrect
R² Linear = 0.142

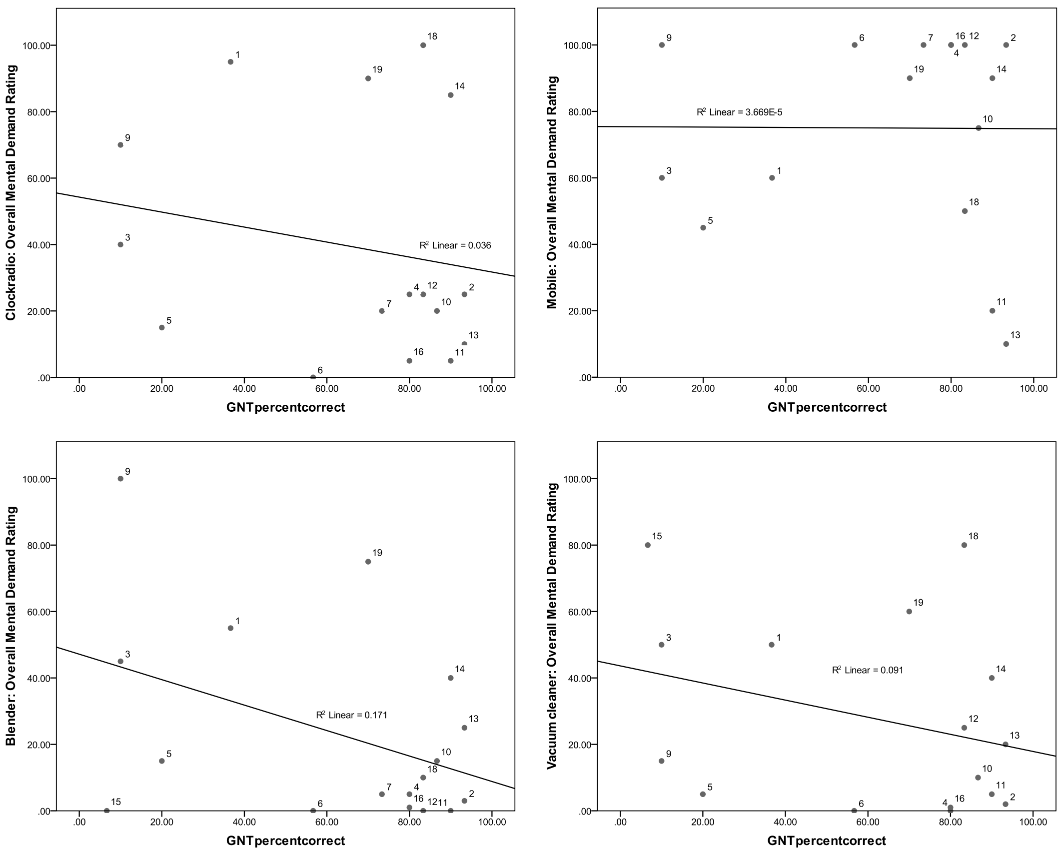
Scatter plots of difficulty in working out next action versus reaction time and GNTpercentcorrect





Scatter plots of overall mental demand versus digit span and span length





Scatter plots of overall mental demand versus reaction time and GNTpercentcorrect

Appendix 12: Multiple Regression Models

Regression Models for Task Time

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.360[a]	.130	-.187	63.46773	.130	.410	4	11	.798	1.653

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: CTaskTimeSec

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-54.981	229.939		-.239	.815	-561.072	451.111					
	DigitSpan	23.681	22.446	.642	1.055	.314	-25.723	73.084	.198	.303	.297	.214	4.683
	SSPspanlength	-3.340	26.866	-.075	-.124	.903	-62.472	55.792	-.103	-.037	-.035	.217	4.614
	RTIfivechoicereactiontime	.194	.331	.220	.587	.569	-.534	.922	-.012	.174	.165	.561	1.781
	GNTpercentcorrect	-.682	1.627	-.351	-.419	.683	-4.263	2.900	.006	-.125	-.118	.113	8.856

a. Dependent Variable: CTaskTimeSec

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.502[a]	.252	-.020	100.57168	.252	.927	4	11	.483	1.422

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: MTaskTimeSec

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	760.569	364.363		2.087	.061	-41.389	1562.528					
	DigitSpan	-44.441	35.568	-.705	-1.249	.237	-122.726	33.845	.015	-.353	-.326	.214	4.683
	SSPspanlength	-28.269	42.572	-.372	-.664	.520	-121.970	65.433	.030	-.196	-.173	.217	4.614
	RTIfivechoicereactiontime	-1.002	.524	-.666	-1.913	.082	-2.156	.151	-.369	-.500	-.499	.561	1.781
	GNTpercentcorrect	2.214	2.579	.666	.858	.409	-3.462	7.889	.043	.251	.224	.113	8.856

a. Dependent Variable: MTaskTimeSec

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.591[a]	.350	.133	71.52871	.350	1.613	4	12	.234	1.512

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: BTaskTimeSec

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	348.903	115.811		3.013	.011	96.573	601.234					
	DigitSpan	2.678	14.954	.054	.179	.861	-29.903	35.259	-.044	.052	.042	.586	1.707
	SSPspanlength	-54.035	25.837	-.905	-2.091	.058	-110.330	2.259	-.397	-.517	-.487	.290	3.454
	RTIfivechoicereactiontime	-.120	.193	-.191	-.619	.547	-.541	.302	-.207	-.176	-.144	.571	1.750
	GNTpercentcorrect	1.225	1.333	.517	.919	.376	-1.680	4.130	-.090	.256	.214	.171	5.834

a. Dependent Variable: BTaskTimeSec

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.752[a]	.566	.408	37.62990	.566	3.583	4	11	.042	1.343

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: VTaskTimeSec

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	115.517	60.928		1.896	.085	-18.585	249.619					
	DigitSpan	17.452	7.878	.575	2.215	.049	.112	34.791	.148	.555	.440	.585	1.709
	SSPspanlength	-14.007	13.659	-.380	-1.025	.327	-44.071	16.058	-.481	-.295	-.204	.287	3.484
	RTIfivechoicereactiontime	.100	.102	.258	.973	.351	-.126	.325	.495	.282	.193	.563	1.775
	GNTpercentcorrect	-.527	.710	-.360	-.743	.473	-2.089	1.034	-.474	-.219	-.148	.169	5.925

a. Dependent Variable: VTaskTimeSec

Regression Models for Errors

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.625[a]	.391	.169	.815	.391	1.763	4	11	.207	1.800

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: CErrors

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	5.897	2.954		1.996	.071	-.604	12.399					
	DigitSpan	-.276	.288	-.488	-.958	.358	-.911	.358	-.448	-.278	-.226	.214	4.683
	SSPspanlength	.036	.345	.052	.104	.919	-.724	.795	-.341	.031	.024	.217	4.614
	RTIfivechoicereactiontime	-.006	.004	-.450	-1.431	.180	-.015	.003	-.055	-.396	-.337	.561	1.781
	GNTpercentcorrect	-.010	.021	-.328	-.469	.648	-.056	.036	-.502	-.140	-.110	.113	8.856

a. Dependent Variable: CErrors

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.351[a]	.123	-.195	2.879	.123	.387	4	11	.814	1.567

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: MErrors

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	10.995	10.430		1.054	.314	-11.962	33.951					
	DigitSpan	-.676	1.018	-.406	-.664	.520	-2.917	1.565	-.112	-.196	-.187	.214	4.683
	SSPspanlength	.363	1.219	.181	.298	.771	-2.319	3.045	.058	.089	.084	.217	4.614
	RTIfivechoicereactiontime	-.015	.015	-.372	-.989	.344	-.048	.018	-.176	-.286	-.279	.561	1.781
	GNTpercentcorrect	-.003	.074	-.031	-.037	.971	-.165	.160	-.059	-.011	-.010	.113	8.856

a. Dependent Variable: MErrors

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.839[a]	.704	.605	.936	.704	7.134	4	12	.004	1.250

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: BErrors

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	6.968	1.516		4.597	.001	3.665	10.271					
	DigitSpan	-.196	.196	-.205	-1.001	.337	-.622	.230	-.505	-.278	-.157	.586	1.707
	SSPspanlength	-1.048	.338	-.905	-3.099	.009	-1.785	-.311	-.819	-.667	-.487	.290	3.454
	RTIfivechoicereactiontime	.000	.003	-.022	-.104	.919	-.006	.005	.162	-.030	-.016	.571	1.750
	GNTpercentcorrect	.010	.017	.215	.567	.581	-.028	.048	-.638	.162	.089	.171	5.834

a. Dependent Variable: BErrors

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.786[a]	.617	.478	1.017	.617	4.439	4	11	.022	2.040

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: VErrors

Coefficients[a]

Model		Unstandardized Coefficients: B	Std. Error	Standardized Coefficients: Beta	t	Sig.	95.0% Confidence Interval for B: Lower Bound	Upper Bound	Correlations: Zero-order	Partial	Part	Collinearity Statistics: Tolerance	VIF
1	(Constant)	3.943	1.647		2.394	.036	.318	7.568					
	DigitSpan	-.154	.213	-.176	-.721	.486	-.622	.315	-.543	-.213	-.135	.585	1.709
	SSPspanlength	-.294	.369	-.277	-.797	.442	-1.107	.518	-.700	-.234	-.149	.287	3.484
	RTIfivechoicereactiontime	.001	.003	.101	.405	.693	-.005	.007	.416	.121	.076	.563	1.775
	GNTpercentcorrect	-.016	.019	-.375	-.826	.426	-.058	.026	-.763	-.242	-.154	.169	5.925

a. Dependent Variable: VErrors

Regression Models for Starting Task

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.583[a]	.340	.101	33.03541	.340	1.419	4	11	.291	1.122

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: CDiffStartTask

Coefficients[a]

Model		Unstandardized Coefficients: B	Std. Error	Standardized Coefficients: Beta	t	Sig.	95.0% Confidence Interval for B: Lower Bound	Upper Bound	Correlations: Zero-order	Partial	Part	Collinearity Statistics: Tolerance	VIF
1	(Constant)	138.943	119.685		1.161	.270	-124.481	402.368					
	DigitSpan	7.859	11.683	.356	.673	.515	-17.856	33.574	.203	.199	.165	.214	4.683
	SSPspanlength	-5.924	13.984	-.223	-.424	.680	-36.703	24.854	-.212	-.127	-.104	.217	4.614
	RTIfivechoicereactiontime	-.218	.172	-.414	-1.267	.231	-.597	.161	-.425	-.357	-.310	.561	1.781
	GNTpercentcorrect	-.422	.847	-.363	-.499	.628	-2.287	1.442	-.124	-.149	-.122	.113	8.856

a. Dependent Variable: CDiffStartTask

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.475[a]	.225	-.056	31.27743	.225	.801	4	11	.549	2.668

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: MDiffStartTask

Coefficients[a]

Model		Unstandardized Coefficients: B	Std. Error	Standardized Coefficients: Beta	t	Sig.	95.0% Confidence Interval for B: Lower Bound	Upper Bound	Correlations: Zero-order	Partial	Part	Collinearity Statistics: Tolerance	VIF
1	(Constant)	141.947	113.316		1.253	.236	-107.459	391.353					
	DigitSpan	-5.531	11.062	-.287	-.500	.627	-29.877	18.816	.159	-.149	-.133	.214	4.683
	SSPspanlength	7.940	13.240	.342	.600	.561	-21.201	37.081	.341	.178	.159	.217	4.614
	RTIfivechoicereactiontime	-.191	.163	-.415	-1.171	.266	-.550	.168	-.360	-.333	-.311	.561	1.781
	GNTpercentcorrect	.023	.802	.023	.029	.977	-1.742	1.788	.239	.009	.008	.113	8.856

a. Dependent Variable: MDiffStartTask

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.657[a]	.432	.243	26.90496	.432	2.284	4	12	.120	1.935

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: BDiffStartTask

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	134.168	43.561		3.080	.010	39.255	229.080					
	DigitSpan	3.430	5.625	.173	.610	.553	-8.826	15.685	-.199	.173	.133	.586	1.707
	SSPspanlength	-6.749	9.719	-.281	-.694	.501	-27.924	14.426	-.455	-.197	-.151	.290	3.454
	RTIfivechoicereactiontime	-.157	.073	-.619	-2.152	.052	-.315	.002	-.260	-.528	-.468	.571	1.750
	GNTpercentcorrect	-.517	.502	-.541	-1.030	.323	-1.609	.576	-.329	-.285	-.224	.171	5.834

a. Dependent Variable: BDiffStartTask

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.632[a]	.399	.180	28.84779	.399	1.825	4	11	.194	1.934

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: VDiffStartTask

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	37.856	46.709		.810	.435	-64.949	140.661					
	DigitSpan	11.488	6.039	.581	1.902	.084	-1.805	24.781	.131	.498	.445	.585	1.709
	SSPspanlength	-4.758	10.472	-.198	-.454	.658	-27.806	18.289	-.413	-.136	-.106	.287	3.484
	RTIfivechoicereactiontime	-.007	.079	-.029	-.094	.926	-.180	.166	.290	-.028	-.022	.563	1.775
	GNTpercentcorrect	-.581	.544	-.608	-1.068	.308	-1.778	.616	-.410	-.307	-.250	.169	5.925

a. Dependent Variable: VDiffStartTask

Regression Models for Selecting Next Action

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.502[a]	.252	-.019	39.35813	.252	.929	4	11	.482	2.122

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: CDiffNextAction

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	185.751	142.591		1.303	.219	-128.091	499.592					
	DigitSpan	2.087	13.919	.085	.150	.884	-28.550	32.723	.020	.045	.039	.214	4.683
	SSPspanlength	-6.872	16.660	-.231	-.412	.688	-43.542	29.797	-.266	-.123	-.108	.217	4.614
	RTIfivechoicereactiontime	-.250	.205	-.425	-1.220	.248	-.702	.201	-.311	-.345	-.318	.561	1.781
	GNTpercentcorrect	-.328	1.009	-.252	-.325	.751	-2.549	1.893	-.228	-.098	-.085	.113	8.856

a. Dependent Variable: CDiffNextAction

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.602[a]	.362	.130	27.62558	.362	1.562	4	11	.252	2.224

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: MDiffNextAction

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	182.304	100.085		1.821	.096	-37.982	402.590					
	DigitSpan	-5.922	9.770	-.316	-.606	.557	-27.426	15.582	.165	-.180	-.146	.214	4.683
	SSPspanlength	9.603	11.694	.425	.821	.429	-16.136	35.341	.339	.240	.198	.217	4.614
	RTIfivechoicereactiontime	-.267	.144	-.597	-1.858	.090	-.584	.049	-.496	-.489	-.447	.561	1.781
	GNTpercentcorrect	-.126	.708	-.127	-.178	.862	-1.685	1.433	.203	-.054	-.043	.113	8.856

a. Dependent Variable: MDiffNextAction

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.691[a]	.477	.303	23.29642	.477	2.735	4	12	.079	1.851

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: BDiffNextAction

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	136.923	37.719		3.630	.003	54.741	219.106					
	DigitSpan	-4.454	4.870	-.249	-.915	.378	-15.066	6.157	-.452	-.255	-.191	.586	1.707
	SSPspanlength	-12.736	8.415	-.587	-1.513	.156	-31.071	5.599	-.557	-.400	-.316	.290	3.454
	RTIfivechoicereactiontime	-.079	.063	-.347	-1.257	.233	-.216	.058	-.165	-.341	-.262	.571	1.750
	GNTpercentcorrect	.045	.434	.052	.104	.919	-.901	.991	-.387	.030	.022	.171	5.834

a. Dependent Variable: BDiffNextAction

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.659[a]	.434	.228	23.47368	.434	2.110	4	11	.148	1.624

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: VDiffNextAction

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	48.503	38.007		1.276	.228	-35.150	132.157					
	DigitSpan	10.741	4.914	.648	2.186	.051	-.075	21.558	.103	.550	.496	.585	1.709
	SSPspanlength	4.000	8.521	.199	.469	.648	-14.754	22.755	-.325	.140	.106	.287	3.484
	RTIfivechoicereactiontime	-.101	.064	-.477	-1.578	.143	-.242	.040	.033	-.430	-.358	.563	1.775
	GNTpercentcorrect	-.952	.443	-1.188	-2.152	.054	-1.927	.022	-.377	-.544	-.488	.169	5.925

a. Dependent Variable: VDiffNextAction

Regression Models for Overall Mental Demand

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.518[a]	.269	.003	35.62855	.269	1.010	4	11	.443	1.712

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: CDiffMentalDemandOverall

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	145.471	129.079		1.127	.284	-138.631	429.572					
	DigitSpan	5.226	12.600	.231	.415	.686	-22.507	32.960	.097	.124	.107	.214	4.683
	SSPspanlength	-7.470	15.082	-.274	-.495	.630	-40.665	25.724	-.266	-.148	-.128	.217	4.614
	RTIfivechoicereactiontime	-.201	.186	-.373	-1.085	.301	-.610	.207	-.326	-.311	-.280	.561	1.781
	GNTpercentcorrect	-.325	.914	-.273	-.356	.729	-2.336	1.686	-.190	-.107	-.092	.113	8.856

a. Dependent Variable: CDiffMentalDemandOverall

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.394[a]	.155	-.152	32.84713	.155	.506	4	11	.733	2.952

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, SSPspanlength, DigitSpan

b. Dependent Variable: MDiffMentalDemandOverall

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	148.193	119.003		1.245	.239	-113.730	410.116					
	DigitSpan	-2.927	11.617	-.151	-.252	.806	-28.495	22.642	.021	-.076	-.070	.214	4.683
	SSPspanlength	7.408	13.904	.317	.533	.605	-23.195	38.011	.108	.159	.148	.217	4.614
	RTIfivechoicereactiontime	-.187	.171	-.403	-1.090	.299	-.563	.190	-.303	-.312	-.302	.561	1.781
	GNTpercentcorrect	-.310	.842	-.303	-.368	.720	-2.163	1.544	-.006	-.110	-.102	.113	8.856

a. Dependent Variable: MDiffMentalDemandOverall

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.732[a]	.535	.381	23.59306	.535	3.458	4	12	.042	1.751

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: BDiffMentalDemandOverall

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	150.402	38.199		3.937	.002	67.173	233.631					
	DigitSpan	.317	4.932	.017	.064	.950	-10.430	11.063	-.334	.019	.013	.586	1.707
	SSPspanlength	-11.631	8.522	-.499	-1.365	.197	-30.199	6.937	-.583	-.367	-.269	.290	3.454
	RTIfivechoicereactiontime	-.133	.064	-.543	-2.087	.059	-.272	.006	-.223	-.516	-.411	.571	1.750
	GNTpercentcorrect	-.289	.440	-.312	-.657	.524	-1.247	.669	-.414	-.186	-.129	.171	5.834

a. Dependent Variable: BDiffMentalDemandOverall

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.661[a]	.437	.232	24.93144	.437	2.133	4	11	.145	1.551

a. Predictors: (Constant), GNTpercentcorrect, RTIfivechoicereactiontime, DigitSpan, SSPspanlength

b. Dependent Variable: VDiffMentalDemandOverall

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-3.448	40.368		-.085	.933	-92.296	85.401					
	DigitSpan	12.710	5.220	.720	2.435	.033	1.222	24.198	.293	.592	.551	.585	1.709
	SSPspanlength	-1.971	9.050	-.092	-.218	.832	-21.890	17.948	-.284	-.066	-.049	.287	3.484
	RTIfivechoicereactiontime	.004	.068	.017	.055	.957	-.146	.153	.304	.017	.012	.563	1.775
	GNTpercentcorrect	-.549	.470	-.644	-1.169	.267	-1.584	.485	-.302	-.332	-.264	.169	5.925

a. Dependent Variable: VDiffMentalDemandOverall