

---

# Design and delivery of a national pilot survey of capabilities

---

Raji Tenneti\*, Joy Goodman-Deane, Patrick Langdon, Sam Waller, Kai Ruggeri, P John Clarkson, Felicia A. Huppert

## Raji Tenneti\*

School of Primary, Aboriginal and Rural Health Care, Faculty of Medicine, Dentistry and Health Sciences,  
University of Western Australia, M706, 35 Stirling Highway, Crawley WA 6009, Australia  
Email: [raji.tenneti@gmail.com](mailto:raji.tenneti@gmail.com)

\* Corresponding author.

## Joy Goodman-Deane, Patrick Langdon, Sam Waller, Kai Ruggeri and P John Clarkson

Engineering Design Centre, Department of Engineering,  
University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK  
Email: [jag76@cam.ac.uk](mailto:jag76@cam.ac.uk)  
Email: [pml24@eng.cam.ac.uk](mailto:pml24@eng.cam.ac.uk)  
Email: [sam.waller@eng.cam.ac.uk](mailto:sam.waller@eng.cam.ac.uk)  
Email: [dar56@cam.ac.uk](mailto:dar56@cam.ac.uk)  
Email: [pjc10@eng.cam.ac.uk](mailto:pjc10@eng.cam.ac.uk)

## Felicia A. Huppert

Well-being Institute, Department of Psychiatry,  
University of Cambridge, Hills Road, Cambridge CB2 0QQ, UK  
Email: [fah2@cam.ac.uk](mailto:fah2@cam.ac.uk)

**Abstract:** Understanding the numbers of people with different levels of ability in the population is important for informing design decisions for mainstream products, but a survey dataset for this purpose does not exist. This paper describes a key step towards obtaining such data. It describes a pilot survey of 362 people from across England and Wales in preparation for a full national survey. The intention of this pilot survey was to demonstrate proof of concept in terms of feasible measures to obtain by a team of interviewers visiting residences. Information was gathered on vision, hearing, hand and arm function, mobility, cognitive function, product use, psychological characteristics, anthropometrics and demographics. An interesting finding is that of those participants who reported any limitations in daily activities due to capability loss, 44% reported limitations due to loss of more than one capability. This finding highlights the importance of measuring multiple capabilities in a single survey. Top-level lessons learnt focused on the feasibility of tests; sampling approach and interviewers' training. They include: simplifying the vision tests; reducing the exclusion criteria for some of the tests; adopting a stratified sampling approach; and allocating more training for interviewers. It is also worth considering using a single touchscreen device to measure multiple capabilities.

**Keywords:** inclusive design; design practice; design tools; user centred design; human factors

**Biographical notes:** Raji Tenneti is a Senior Research Officer at the University of Western Australia (WA), where she primarily works on health data to evaluate and inform clinical service provision in the WA health system. Previous research positions were in the UK at University of Cambridge, where this work was completed, and at University of Strathclyde. Raji holds a PhD in Computer-Aided Design. She has diverse research interests which include engineering product design, human-computer interaction, psychology, user experience research, public health and health service utilisations along with statistics. She has been

involved in analysing health-related datasets of people in the UK and Australia. Raji was one of the main leads in designing, executing and analysing the Towards Better Design Survey (2010) conducted across England and Wales, which is publicly available at <http://discover.ukdataservice.ac.uk/catalogue?sn=6997>

Joy Goodman-Deane is a Research Associate at the Engineering Design Centre, University of Cambridge. Her research interests focus on inclusive design, that is the design of mainstream products and services that are usable by as many people as possible. Recent work examines how to support designers in designing inclusively. This includes the production of design tools such as software and wearable devices that simulate capability loss. Another tool is the exclusion calculator that estimates the proportion of people who would be excluded from using a product because its demands on their capabilities are too high. Joy's interests also include examining how capabilities and technology experience vary across the population, and investigating the use of design tools and user data in design practice.

Patrick Langdon is a Senior Research Associate at the Engineering Design Centre, University of Cambridge. He was employed by an Artificial Intelligence Vision Research Unit at the University of Sheffield working on robotic vision. Subsequently, at Lancaster University EDC, he worked on the application of AI and user centred design principles to complex engineering design systems. At Cambridge his most recent research is in “inclusive design” and multimodal interface for interactive digital TV that will provide an adaptable input and output for impaired users based on profiling from advanced cognitive modelling. He has recently joined phase 2 of the India-UK Advanced Technology Centre, that focuses on the state-of-the-art research into interface design for ageing and impairment to be transferred into practical applications in rural and urban contexts, having impact in both the UK and India.

Sam Waller is a Research Associate within the Cambridge Engineering Design Centre at the University of Cambridge. Sam produced the inclusive design toolkit, which is freely available at [www.inclusivedesigntoolkit.com](http://www.inclusivedesigntoolkit.com), which includes design tools such as an exclusion calculator and impairment simulator. Sam's research interests include analysing population statistics and creating design tools that enable practising designers to quantify the number of people who would not be able to do a particular task, based on the demands placed on the user's abilities. Sam led the EDC's contribution to designing and analysing the Towards Better Design Survey (2010), which is publicly available at <http://discover.ukdataservice.ac.uk/catalogue?sn=6997>

Kai Ruggeri is Research Associate in Health Policy at the University of Cambridge, Engineering Design Centre, and Affiliated Lecturer in the Department of Psychology. He leads a variety of quantitative projects primarily in the area of global health policy. Previous research positions include the Cambridge Institute of Public Health and the Department of Psychiatry, where this work was completed. He also serves as the Director of the Junior Researcher Programme for psychology students in Europe and an economic consultant in the US and UK healthcare systems as well as Visiting Faculty in Microeconomics at the HEC Paris' Executive MBA programme in Qatar. Kai holds a PhD in psychology from Queen's University in Belfast.

P John Clarkson is Director of the Engineering Design Centre, University of Cambridge. Prior to this, he worked as a Manager of the Advanced Process Group at PA Consulting Group's Technology Division. At PA, John gained wide experience of product development with a particular focus on the design of medical equipment and high-integrity systems, where clients required a risk-based systems approach to design to ensure timely delivery of safe systems. His research interests are in the general area of engineering design, particularly the development of design methodologies to address specific design issues, for example, process management, change management, healthcare design and inclusive design. As well as publishing over 500 papers, he has written and edited a number of books on medical equipment design, inclusive design and process management.

Felicia A. Huppert is Emeritus Professor of Psychology and the founding Director of the Well-being Institute at the University of Cambridge. Felicia also has a part-time professorship in the Centre for Positive Psychology and Education at the University of Western Sydney. She

has been involved in the design and analysis of major longitudinal and population studies across the life course, and she headed the consortium that has developed national indicators of well-being for Europe as part of the European Social Survey. She is currently a member of the Technical Advisory Group of the UK Office for National Statistics 'Measuring National Well-being Programme', as well as an advisor on similar European and OECD initiatives.

---

## 1 Introduction

Good design is design which is well-matched to the users' characteristics and capabilities. To help produce such designs, good data on those characteristics and capabilities is needed. Although there are some datasets available, many of these are based on small samples, while the population-level surveys usually focus on one specific aspect of capability. In contrast, product use engages several abilities in combination. For example, using a mobile phone successfully requires a certain level of vision and a certain level of dexterity, among other capabilities. Thus it is important to know not just how vision varies in the population and separately how dexterity varies. It is also important to know how many people have the particular combinations of vision and dexterity abilities required to use particular products.

The latest inclusive design research program (i~design, 2009) aimed to address this shortcoming directly, by designing and undertaking a national survey specifically for this purpose. The aim was to collect new, up-to-date information which is representative of the population on the capabilities that are particularly relevant to design. This paper describes a key step towards this, in the form of a pilot survey of 362 people drawn from across England and Wales. This national pilot tested the survey methods and materials in preparation for the full national survey. The work presented in this paper describes the design, execution and analysis of the pilot study, before discussing the implications for the full national survey.

There is an increasing awareness of the need to design for the widest possible range of users. This is typically focused on the needs of older and disabled people, to meet the growing body of disability discrimination legislation (Goodman et al, 2006). However, it also has implications for all users. Many people find products and services difficult to use. Some find them frustrating, some struggle but can manage if necessary, and others are effectively excluded from using them altogether (Clarkson et al, 2007; Waller et al, 2009a). There are many reasons for these difficulties, but a root cause is that many products and services do not match the users' characteristics, capabilities and ways of thinking (Keates and Clarkson, 2003).

Inclusive design seeks to address this by making mainstream products usable by as many people as reasonably possible, without requiring them to use specialised adaptations (Clarkson et al, 2007). To do this, designers need good information on users' characteristics, such as databases of users' capabilities for the design of environments and products (McGinley et al, 2010). They also need tools that present the relevant data in an accessible and useful way (Johnson et al, 2010).

One such tool with great potential for assisting the implementation of inclusive design is the Exclusion Audit (Waller et al, 2009b). In this method, the designer or usability expert identifies the demands a product or service places on the user's capabilities. These demands are then compared with data about the population to estimate how many people would be excluded from using the product because their capabilities do not meet these demands (Goodman-Deane et al, 2011). This can help to assess how inclusive the product is, and is useful for comparing designs and identifying points for improvement.

An exclusion audit of this type is dependent on good population-level data about users' capabilities. Furthermore, product use involves several capabilities in combination. To determine exclusion, it is thus important to know how many people have the particular combinations of capabilities required to use particular products. This requires data on a range of capabilities for each individual in the sample. Such data is also important when considering accessibility. For example, a common solution for vision impairment is to provide supplementary audio information, yet the success of this strategy will depend on the extent to which people who have impaired vision still have

full hearing. An examination of data from the Disability Follow-Up Survey (DFS) (Waller et al, 2010a) indicates that the majority of people who cannot perform tasks associated with one capability are also unable to perform tasks associated with another type of capability. This indicates a critical need for treatment of co-occurrence of impairment within inclusive and accessible design.

### *1.1 Existing datasets*

There are several existing datasets of users' capabilities. However, each has its own limitations in giving detailed data on individuals for the purpose of inclusive design. For example, HADRIAN (Porter et al, 2004) covers a range of capabilities but it has a sample of 100 participants only, deliberately biased towards older and disabled participants. The Geron 1998 Dutch Elderly study was biased due to the high educational level of the sample (Steenbekkers and van Beijsterveldt, 1998). ADAPS (Molenbroek, 1987), a computer-aided design model, uses a wider population of 822 elderly Dutch people, but focuses only on functional body dimensions thereby limiting the data to a few specific aspects of capability. Ergonomics databases such as Adultdata (Peebles and Norris, 1998), Older Adultdata (Smith et al, 2000) and Childata (Norris and Wilson, 1995) gather together a range of capability measurements. However, these draw the data for each capability from separate sources, thereby making it impossible to say how the different capabilities interact.

In fact, Johnson et al (2010) found that no currently available dataset meets the necessary criteria for exclusion audits, and the best available dataset was the ONS 1996/97 DFS (Grundy et al, 1999). This survey contained detailed questions on 13 ability categories, seven of which are most relevant to product use (Waller et al, 2010b). However, it was not designed for assessing products but to provide a measure of disability severity. Thus it does not cover all the capabilities involved in product use and does not examine them in sufficient detail. Furthermore, it suffered from an untested theoretical basis and sampling biases, and was based only on self-report data (Waller et al, 2010a).

Therefore, a new dataset is required, providing up-to-date, population-level data on a whole range of capabilities of relevance to design. This paper describes a key step towards obtaining such data. It describes a pilot survey of 362 people from across England and Wales in preparation for a full national survey.

### *1.2 Preparatory study*

In preparation for the survey, initial theoretical research reviewed existing databases and examined broader issues around measuring capability. Johnson et al (2010) reviewed the potential influences on measures of capability and concluded that a number of issues need to be considered for the construction of a survey to reliably assess capabilities. Specific issues include: self-report versus performance measures; granularity of measurement; psychological characteristics; and naturalistic versus experimental settings for performance.

The theoretical research was followed by an exploratory study with 100 people (Tenneti et al, 2012). Its objective was to establish the extent to which component functions (i.e. simple sensory, motor and cognitive capabilities), performance on component activities (i.e. smaller tasks within the larger product interaction) and psychological characteristics could predict human-product interactions and hence be used as proxies in the future national survey. The study involved participants completing self-report and performance-based measures. Results showed that performance measures of component functions, self-report and performance measures of component activities and a very few self-reported psychological characteristics are strong predictors of successful product interactions. Based on the results of the exploratory study, some items were retained in the national pilot survey on the basis that they are likely to be significant predictors of product interaction. In addition, the results strongly suggested that self-report and performance measures may be measuring different aspects of functioning, and thus produce different but valuable information. Collectively, the findings from this stage informed the choice of measures and ensured that instructions, materials, tests, interview questions and response categories were set at the most appropriate levels for the national pilot survey.

## 2 Method

The aim of this pilot survey was to gather data on a range of user capabilities and characteristics in preparation for a larger, national survey. The survey thus tested the feasibility of the methods and provided suggestions for changes to improve their usefulness in future surveys. Furthermore, the survey examined the extent to which people have multiple capability losses. It also examined the spread of measurements within each capability, to examine whether or not the tests used cover the full range of capabilities in the sample.

Ethical approval for the national pilot survey was granted from the Cambridge University Psychology Research Ethics Committee (Application No: 2010.54). The survey itself was conducted by the National Centre for Social Research, a non-commercial professional survey organisation, under the direction of the authors.

Before conducting the survey, a smaller pilot was undertaken with 16 participants to test the comprehension of specific interview modules, to time these modules and to test the accuracy of the interview program. The pre-pilot found that the basic format of the interview and most items worked well, but recommended that the interview length be reduced to an hour to maximise response. As a result, less standard items such as symbol test from cognitive function module were removed (see Section 2.3.5), buying preferences and technology experience items in product use module were reduced (see Section 2.3.6), and anthropometric information, especially height and weight data were collected by self-report rather than being actually measured (see Section 2.3.8).

### 2.1 Data archive

Further details on the survey and all final data, including derived and aggregate variables, weights and documentation are stored in the UK Data Archive (Clarkson et al, 2012). This allows long-term preservation of the data and its distribution to other groups who are interested to investigate or develop models of various capabilities.

### 2.2 Participants

The sample was designed to be representative of the general population in England and Wales aged 16+ and living in private households. A sample of 990 postcode addresses was drawn from 30 primary sampling units within England and Wales. Households not containing anyone in that age range were not eligible for the survey. At responding households interviewers selected one individual aged 16+ at random. The response rate was 37% of the issued sample or 40% of the eligible sample.

### 2.3 Procedure

The survey was administered face-to-face in participants' homes so that the testing environment would be similar to that in which products are typically used. It was also more practical than an assessment centre for a survey spread throughout the country. A computer assisted personal interviewing (CAPI) programme on a laptop was used to guide the interviewer and allow input of answers.

The survey comprised a series of modules, covering the main domains of human capabilities that impact on product use: vision, hearing, hand and arm function, mobility and cognitive function. There were also modules examining product use, psychological characteristics, anthropometrics and socio-demographic information, as these also affect product use. Both self-report questions and performance tests were used in the above modules to assess participants' capabilities.

Where possible, the questions and tests used were based on well-established validated methods. In some cases, the tests had to be adapted to fit the practical and time constraints of the survey, but such adaptation was kept to a minimum. Care was taken to include only those items which provided the best balance between time, cost, accuracy and relevance to product use. The items are summarised below. The full survey questionnaire is freely available within the technical report in the UK data archive (Clarkson et al, 2012).

### 2.3.1 Vision

Vision was measured as it is key for interacting with any product with visual information, including packaging, instructions, and electronic devices with screens (e.g. mobile phones).

Self-report questions adapted from NIH Toolbox (Gershon et al, 2010) covered specific vision conditions, the use of vision aids, and limitations in daily activities due to vision. In addition, participants were given a series of performance tests. The intention of conducting these tests in participants' homes was to measure vision in a real product use environment, with the lighting levels and other conditions that users are likely to encounter in practice when using products (Elton et al, 2013).

The vision tests were conducted using LogMAR vision charts (Hazel and Elliotts, 2002) with rows of block letters at decreasing sizes. It was intended to capture the comfort vision of participants, hence a protocol was designed where the participants were asked to start reading on the smallest row they found comfortable to read. If they read their "comfort row" successfully, they then read down the chart (smaller letters) until they failed to read a row or they gave up. Success on a row was defined as making one or fewer errors on it. If they did not read their "comfort row" successfully, they read up the chart (larger letters) until they did read a row successfully. This identified the smallest row that participants could read successfully.

Three vision charts were used: (i) a near vision chart with high contrast (70%) letters, (ii) a near vision chart with low contrast (30%) letters, and (iii) a distance vision chart with very high contrast (100%) letters. Participants were asked to hold the near vision charts at a comfortable reading distance, as this corresponds to what happens in practice when using products. The distance vision tests were conducted at 3m. All the tests were done with (i) the vision aids (if any) that the participants used for the majority of the day and (ii) the vision aids (if different, and if they were available) used for close or distance vision, as appropriate.

### 2.3.2 Hearing

Measures of hearing were included as many products use audio output, such as speech, ringers and alarms.

Self-report questions, including use of hearing aids, limitation in daily activities due to hearing and specific hearing capabilities, were adapted from the National Study of Hearing (Davis, 1995). In addition, a hearing test to help inform the design of audio in products, such as speech, ringtones and auditory feedback was performed.

Standard pure tone threshold tests were not possible because of the uncontrolled background noise in participants' homes. Speech and speech-in-noise tests were used instead. These are less sensitive and thus less affected by background noise. The tests were based on the Hearing Check used by the RNID (RNID, 2013) but were simplified due to time and equipment restrictions. The simplifications mean that the hearing measures are not so precise, but can still provide useful data, as they indicate broad levels of hearing in actual product use situations. The tests used an MP3 player (Sandisk Sansa Clip) and over-ear headphones (Audio Technica ATH-PRO5 V). Where used, hearing aids were removed for the tasks, to avoid interference with the headphones.

A series of sound clips was played. Each contained three numerical digits, which participants repeated aloud. A small number of digits were used to ensure that the test measured hearing, not memory. Participants were played two clips in each condition to give a score out of six. Firstly, the digits were played at three volumes (60dB, 40dB and 25dB) without any background noise. The lowest volume at which participants could correctly identify five of the six digits was recorded as the measure of hearing without noise. Digits were then played at medium volume (40dB) with one of three levels of background noise taken from the ICRA noise CD (Dreschler et al, 2001): easy (25dB), medium (40dB) and difficult (60dB). The highest level of background noise at which participants could correctly identify five digits was used as the measure of hearing with noise.

### 2.3.3 *Hand and arm function*

Hand and arm function was measured because this capability is important for interacting with products that require manual manipulation, such as pressing buttons, manipulating switches, picking up and carrying objects and inserting cables. The assessment of hand and arm function covered various capabilities of the upper body using self-report questions and performance tests.

Participants were first asked self-report questions about their difficulty in performing specific everyday tasks with their hands, such as opening a bottle with a plastic screw cap and tying a shoe lace. These tasks were chosen to reflect distinct types of hand function commonly used in interacting with products. For example, opening the screw cap requires pinch grip with one hand while using the other hand to stabilise the bottle. Next, the Moberg pickup test measured a combination of pinch grip and fine finger dexterity (Ng et al, 1999). This required participants to pick up 12 small objects from a table and place them in a container as quickly as possible. The test was timed for each hand separately. Faster times indicate greater dexterity. Participants' hand grip strength was then measured using a Smedley-Type Hand Dynamometer (100kg). Measures were taken (where possible) seated with participants' backs upright and unsupported and their forearm supported and at right angles to the body. Two separate measures were taken for each hand: comfort (where participants grip at a comfortable level for two seconds) and threshold (where they grip as hard as they can for two seconds).

This was followed by tests involving simple actions to assess specific hand functions involved in using everyday products. The ability to press different sized buttons was assessed using a mobile telephone. Participants pressed a sequence of four small buttons and then four medium sized buttons. In both cases, the interviewer recorded the screen shown after the task. For the small buttons, this indicated whether the participant had pressed the buttons correctly in the specified order. However, an experimental weakness meant that, for the medium buttons, this only indicated whether the last of the buttons had been pressed correctly. The ability to insert different types of cables was then assessed. Participants were asked to insert a round, tapered headphone jack and then a trapezium-shaped USB cable into an MP3 player. The interviewer scored whether the tasks were completed successfully. For more detail, see Bradley et al (2012).

Finally, participants' reach and stretch ability was assessed by asking participants to stretch both arms out in front and then above shoulder level for a couple of seconds. The interviewer scored the extent to which each arm was extended, choosing from fully, partially or not extended. The idea behind these tasks came from the arm function questions asked in ELSA (Marmot et al, 2003).

### 2.3.4 *Mobility*

Mobility was assessed through self-report questions about mobility aids, getting up from a chair and reaching down, as these actions impact on whether people have hands available to interact with products and can reach product components. The questions were adapted from ELSA (Marmot et al, 2003). General walking ability was not assessed through a performance test because it has limited impact on product use, and to keep the length of the survey down.

### 2.3.5 *Cognitive function*

Cognitive function was measured because it is important for using interactive products. Memory is needed to learn how to operate a product and remember where to find options in menu structures, among other things. Executive function is important for planning actions in product interaction. Furthermore, many products rely on some level of literacy or numeracy, as they provide written or numerical information.

Standard tests from ELSA (Steel et al, 2003; Huppert et al, 2006) were used to assess the following domains of cognitive function of direct relevance to successful product interaction: learning and memory; executive function; and basic skills (literacy and numeracy).

Immediate and delayed memory were assessed using a list of 10 words presented using a pre-recorded human voice from a computer (volume adjusted when required). The participant was asked to recall the words immediately, and again after a short delay during which they performed another

task. The combined measure was the total number of words correctly recalled out of 20. Prospective memory (i.e. memory for future actions) was assessed by asking participants to remember to carry out an instruction later in the session. A correct response required the person to carry out this action without being reminded.

Executive function was assessed using a visual search task. Participants were given a page of letters and were asked to cross out as many target letters (P and W) as possible in one minute. The total number of letters searched gives a measure of processing speed. Search efficiency is estimated based on the number of target letters identified correctly divided by the total number of letters searched.

Basic skills were assessed using literacy and numeracy tasks. Literacy was assessed using a piece of text, followed by three comprehension questions. Similarly, numeracy was assessed using four mental arithmetic questions based on real-life situations. The total number of correct answers for each gives a measure of literacy and numeracy scores respectively.

### 2.3.6 *Product use*

The survey also examined participants' prior experience with technology, and various aspects of how they use products, as this can influence how they would use products in the future. The products chosen for the survey represent a range of common technological interfaces.

Technology experience was assessed by participants rating on a scale of 1 to 4 how often they performed various technology-related tasks, such as making a phone call on a mobile telephone. If they had done a task, they also rated how easy or difficult they found it on a scale of 1 to 6. The experience questions and scales were adapted from (Blackler, 2006) and the ease of use questions and scales from the Philips survey (Philips, 2004). More details can be found in (Bradley et al, 2013).

Participants were then asked how they normally learn to use a new product, and to identify the three factors (from a list of seven possibilities) that they consider most important when buying new products. This question was adapted from the Philips survey (Philips, 2004).

The survey then assessed participants' ability to use two different styles of product interface to select items from a list. In the 'select and confirm' style, users navigate up and down the list using arrow keys and press a 'select' key when they reach their choice. In the 'number navigation' style, users choose an item by entering the number next to it. These styles were chosen as they are both commonly used in digital products, and differ significantly from each other. Previous experiences with users also indicated that these interfaces have the potential to cause significant confusion, particularly among older users. The use of the different interface styles was assessed using a simplified paper prototyping method (Snyder, 2003). Participants were shown images of mobile telephones with these interfaces, and were asked to indicate which buttons they would press to choose a particular menu option. The order in which the interfaces were presented was counter-balanced. The interviewer scored whether the correct buttons were pressed in the correct order. See (Bradley et al, 2012) for more details.

### 2.3.7 *Psychological characteristics*

Product use is affected by people's feelings of self-efficacy (their beliefs about their ability to reach certain levels of performance) and their anxiety about products. Thus the survey included some self-report questions relating to general self-efficacy, product self-efficacy and product-phobia. These were adapted from the General Self-Efficacy Scale (Schwarzer, 1993) and the exploratory study (Tenneti et al, 2012).

### 2.3.8 *Anthropometrics and demographics*

Some anthropometric information of particular relevance to product use was gathered. Height and weight were examined as they give an indication of general body size, influencing what users can reach, e.g. shelves in a supermarket. They were collected using self-report. This was considered adequate because previous surveys (Erens et al, 2001; The Welsh Health Survey, 2003) reported that self-reported height and weight yield appropriate data. Finger size is also relevant to the sizes of



products, e.g. the clearance needed in between buttons. The diameter of the index finger on the participant's dominant hand just past the first finger joint was measured using a finger measurement plate. This has holes cut in it increasing in diameter in 2mm increments. The interviewer recorded the smallest hole through which the participant could insert their finger the required amount. This gives a ballpoint measurement of finger diameter, and provides an indication of hand size that can inform the design of product controls. Information was also collected on socio-demographic characteristics, including age, sex, marital status, living conditions, education, employment, occupation, ethnicity and income. The questions were adapted from ELSA (Marmot et al, 2003). This information was gathered to ensure that the sample was balanced, to help understand the spread of capabilities across the population, and to enable the examination of different market segments.

## **2.4 Analysis**

Most of the performance measures were derived without difficulty. However, some of the vision measures were found to be internally inconsistent. Further examination indicated that several interviewers had recorded the results of the vision threshold tests incorrectly. For example, interviewers frequently recorded the smallest row that participants got correct rather than the largest row that they made errors on. Altogether over half the participants were affected. This was too many to just record as missing data. Instead, the vision data was subjected to significant recoding. By examining the original recording booklets from the interviews, it was possible to determine the correct vision results in most cases. Cases where the data could not be confidently determined were coded as missing values. Section 4.5.1 discusses how these issues could be minimised in the future main survey.

The results were first analysed for distribution, then were either dichotomised or grouped. For example, in analysis of impairment figures, individuals were often categorised as either "passing" (no impairment) or "failing" (impairment) a particular task. This was done to control for outliers and to enable impairment analysis.

For analysing combinations of impairments, missing values were treated as passes (no impairment). That is, if the reason for a missing value was ambiguous, it was assumed that this was not due to impairment. For example, if an item was missing due to lack of consent for an unknown reason, the participant was treated as passing that item. However, if it was missing because the participant had recent surgery and thus could not perform the test, it was considered as a fail. This approach results in conservative estimates of impairment. The actual levels of impairment are likely to be higher than those reported in this paper.

Due to limited space and the focus of this paper, all the results in this paper are presented using descriptive statistics, e.g. means and percentages of participants with different capabilities. Some comparative analyses by age and gender are presented in other related papers (e.g. Bradley et al, 2012; Bradley et al, 2013).

## **3 Results**

There is not enough room here to present the full results of the survey. Rather, this section summarises some of the main statistics in each module, in order to familiarise the reader with the basic results and prepare for the subsequent discussion, which examines (among other issues) data issues such as ceiling and floor effects. The full dataset is available from the UK Data Archive (Clarkson et al, 2012), and fuller results on particular modules are presented in other papers (e.g. Bradley et al, 2012; Bradley et al, 2013; Goodman-Deane et al, 2013).

The results which follow are based on the raw, unweighted data, but sample weights are available in the data file and can be applied if needed. Missing values are excluded from the total when calculating percentages and presenting the results.

### **3.1 Demographics**

Of the 362 total participants, 53.6% were female. Age was spread 16-34 (23.5%), 35-49 (29.3%), 50-64 (23.8%) and over 65 (23.5%). Occupations were grouped as managerial/professional (27.9%),

intermediate (46.4%) and manual/routine (20.7%). Education was classed by no qualification (23.7%), intermediate qualification (55.0%), degree or equivalent (18.4%) or other (2.8%). The vast majority of the sample was white (91.7%). Marital status among participants was split between never married (25.4%), divorced-widowed-separated (21.8%) and married or living with partner (52.8%). Household incomes were 43.8% under £20,000, 32.1% between £20,000 and £39,999, and 24.1% £40,000 or more per year. The median household size was two people, with 30.7% living alone and 69.3% living with one or more others.

### 3.2 *Capability measures*

A summary of the key capability measures is presented in Table 1. In general, the best performance for each measure is shown in the left-most column in the table.

[insert Table 1]

#### 3.2.1 *Vision*

Of the sample, 65.2% reported using a vision aid for any activity (including normal glasses/contact lenses for the majority of the day, and aids specifically for reading/distance-vision).

For the vision tests, the vast majority of the sample was tested during the day (85.2%) with very bright (17.6%), adequate (79.9%) and poor (2.5%) lighting conditions in the households. For distance-vision (see Table 1), with vision aid if available, 69.3% of the participants had a Snellen visual acuity of 6/6 or better, roughly equating to “normal” vision. 3.8% had 6/15 or worse, making it illegal for them to drive in the UK. The remaining 26.8% had visual acuity between 6/7.5 and 6/12, which is generally considered worse than “normal” vision but does not make driving illegal. Results for near-vision are also shown in Table 1. 60.7% of participants could read the smallest or second smallest row (0.6 and 0.8mm), and nearly 90% could successfully read the fourth smallest (1.2mm).

#### 3.2.2 *Hearing*

4.1% of the sample reported using a hearing aid (although 7.7% reported owning one). As shown in Table 1, over 95% of participants could hear sounds at medium volume, even with a medium level of background noise.

#### 3.2.3 *Hand and arm function*

Arthritis in one or both hands was reported by 15.7% of the sample, while 15.7% reported a recent back or shoulder injury or operation.

Performance measures are shown in Table 1. Moberg pickup test results ranged from 7 to 60 seconds (scores higher than 60 were treated as 60) with a median time of 13 seconds. Grip strength results (both threshold and comfort values) were higher for men than women. Successful completion rates were fairly high on most of the everyday dexterity tasks. Nevertheless, over 20% of the sample could not press the small buttons correctly. Furthermore, there was a problem with recording the results from the medium-sized button task, meaning that the actual success rate on this task is likely to be lower than that shown in Table 1.

15.7% of participants were not asked to do the reach and stretch test, due to a recent back or shoulder injury or operation. However, it seems likely that many of these people would experience some reduction in their reach and stretch ability. Of the remainder, 5.9% (5.0% of the total sample) were unable to fully extend one or both arms.

#### 3.2.4 *Mobility*

8.6% of the sample reported using a mobility aid. The responses to the self-report performance questions are shown in Table 1. 18.2% reported some difficulty with getting out of a chair with arms,

and 22.4% with a chair without arms. 24.6% had difficulty reaching to floor level (e.g. to plug something in).

### 3.2.5 *Cognitive function*

As seen in Table 1, the median number of words recalled across immediate and delayed memory was 11, with a range of 1 to 20. Executive function was assessed using a visual search task. The mean number of letters scanned was 309.4 (SD=96.5), and the mean proportion of letters correctly identified was 83.1% (SD=14.0%). Literacy and numeracy results are also shown in the table. Note that 40.6% answered all of the literacy and numeracy questions correctly, indicating that care needs to be taken over assumptions about literacy and numeracy capabilities.

### 3.2.6 *Self-reported limitations*

In each module, participants were asked if their daily activities were limited because of the capabilities described in that module. The responses are summarised in Table 2.

It is worth noting that only 20% of those that reported a hearing limitation use a hearing aid, and that 46% of those who own a hearing aid do not use it. Interestingly, of those that both own and use the aid, only 20% reported any limitation.

[insert Table 2]

## 3.3 *Other measures*

The results for frequency and ease of technology use are shown in Table 3. The percentage of those who used each technology occasionally or frequently varied from 21.1% or 93.6% depending on the technology. Relatively small proportions of those who had used the technologies reported them being difficult to use. More details can be found in Bradley et al (2013).

The main approaches to learning how to use a new product were trying it out and seeing what happens (39.6%), reading the manual (45.4%) and asking to be shown (15.1%).

[insert Table 3]

Table 4 shows the factors considered most important when buying new products. Quality (31.8%), ease of use (22.5%), range of functions (17.5%) and low price (16.1%) were the most popular first considerations.

[insert Table 4]

The survey also assessed participants' ability to use two different styles of product interface. 73.4% of the sample successfully completed the task with the 'select and confirm' interface, while 61.8% were successful with the 'number navigation' interface. See (Bradley et al, 2012) for more details.

Participants were then asked to rate their feelings of self-efficacy and anxiety about products. 74.7% agreed that they could always find a solution to a problem, and 76.6% were confident in their skills and ability to use new products. Conversely, only 19.7% felt like giving up easily when things go wrong, and 42.6% felt nervous with new products.

The height and weight results are shown in Table 5. Finger width was also measured (see Section 2.3.8). Most (91%) participants had finger widths between 13mm and 19mm, with a median of 17mm.

[insert Table 5]

### **3.4 Combinations of capabilities**

One of the aims of this survey was to collect data on a range of capabilities from a single sample. This enables us to examine combinations of capabilities and capability loss. One reason this is important is because designers address capability loss by providing alternative interactions. For example, vision loss may be addressed by providing audio information. However, some people have both vision and hearing loss, and so would not benefit from this design solution. To understand the impact of this, we need to understand how many people have both losses.

There is not room here to examine the combinations of all the capabilities surveyed. This is particularly complex because different capabilities were measured on different scales, and there are different definitions of what counts as “capability loss”. As an overview, we examine reported limitations in daily activities across all the capability categories surveyed (Section 3.2.6). These variables were used because they provide a summary of the categories related to daily life, and were all measured on the same scales.

The prevalence of combinations of capability loss is shown in Table 6. 62.7% of the sample did not report any limitations. Of those that reported did, 43.7% of them had limitations due to more than one capability.

[insert Table 6]

## **4 Implications for future survey**

The intention of the pilot survey was to inform a full national survey of capabilities that are particularly relevant to design. We hoped to identify issues in the conduct of the survey and the collection and analysis of the data that could be improved before the national roll-out. Analysis of pilot data can also inform decisions about which measures to include in the full survey.

### **4.1 Sampling issues**

Some suggestions for improving the survey response rate were given by the interviewers. These include providing an incentive to the participants. Adopting a proportional stratified sampling approach (Van Dalen, 1979) can also help to obtain a more representative sample.

Another sampling issue is that certain capabilities are currently required to complete the survey. For example, vision is required to do some of the cognitive tests, although the cognitive abilities measured are independent of vision. In addition, people with moderate to severe cognitive impairments are essentially excluded from taking part in the survey altogether. The survey thus misses out on critical input from certain groups with special needs. To address this issue, it may be necessary to produce tailored versions of the survey for these groups. The initial contact with these participants may have to be done in a different manner, perhaps through organisations that work with these groups. The interviewers will also have to be more carefully selected and trained to deal with differently abled participants. Instances where more interviewers training is required for testing protocols and operating equipment is discussed in Section 4.5.

### **4.2 Missing data**

Missing data is unavoidable but it is desirable to reduce the amount of it as much as possible. In this pilot survey, wherever possible, missing data was coded to indicate the reason why the data was missing. This best supports future use of the data to plan future surveys.

Sometimes individual participants did not complete an item or module because the test could not be conducted for practical or safety reasons. A particular instance is in the Reach & Stretch module where 57 participants were excluded from the tasks due to back or shoulder injuries. This criteria may have been over-cautious as such injuries could vary from minor to severe, and the tasks themselves were very simple with the participants instructed to stop if the movement was not comfortable. Thus the exclusion criteria could be reduced for the full survey, although care should

still be taken. Another issue was limited space for distance vision tests in some of the participants' homes. This resulted in 27 participants not doing these tests. It may be possible to reduce this distance in the full survey. Technical problems with the testing equipment for hearing also produced 11 missing cases. Thus this equipment should be thoroughly checked or backup devices provided to interviewers.

### **4.3 Data distribution**

#### *4.3.1 Ceiling/floor effects*

Ceiling effects usually happen when a question or test is too easy so that many participants score very highly. Conversely, floor effects occur when a question or a test is too difficult so that most participants score very low. In most items, ceiling and floor effects were not a major issue. However, there were ceiling effects in a few tests. In particular, 83% of the participants could successfully hear the most difficult level on the hearing tests. To better capture the variation in the population, it would be good to begin the hearing tasks at lower decibel levels. In the vision tests, 60.5% could read the smallest or the second smallest row on the near vision chart. While it would be useful to print a smaller row if possible, this may be limited by printer resolution for the near vision charts. In addition, it may not be beneficial to designers to know about letter sizes and population percentages smaller than this.

A majority (59.8%) of the sample made no errors on the literacy test, and a similar number on the numeracy test. While more complex tests are available, they are more time-consuming and may not be so relevant to product use. What is initially needed is a simple measure of who fails basic literacy and numeracy. The Reach & Stretch tests also showed ceiling effects, with 85.5% of the sample successfully completing them (72% if we assume that people with back and shoulder injuries would not successfully complete the tests). However, this may be acceptable because (in conjunction with height measures) it can give a reasonable indicator of the reach needed in product use. There were also high numbers (77-92%) who passed the dexterity product tasks. However, these results can be used in conjunction with results from the Moberg tests to give a fuller picture of dexterity capability.

#### *4.3.2 Granularity of measurements*

For any questionnaire with discrete category outcomes, the number and scope of categories provided should be chosen to best capture the variation in the population sampled, as quickly as possible. The results from this pilot survey can inform the choice of categories provided in future surveys. For example, the results from the vision tests indicated that very few people made errors on the largest rows. This indicates that this level of granularity may not be needed at this end of the vision test, and the number of rows presented could be reduced.

#### *4.3.3 Multiple impairments*

Of those with any significant capability loss, 44% had losses of more than one capability (see Section 3.4). This indicates the importance of measuring multiple capabilities within a single survey. This is relevant when designing any product, service or built environment whose use involves more than one capability. It is important to ensure that the needs of individuals with multiple capability losses are addressed. To identify the numbers and characteristics of those individuals we need a single survey covering all the relevant capabilities.

This also relates to the development of personas with impairments (Fennell et al, 2011). Our survey indicates that these need to include personas with multiple impairments. For example, it is not enough to have one persona with a vision impairment and another with a hearing impairment. There are significant numbers of people who have both vision and hearing impairments. If a product or service is designed to ensure that people who have vision impairments can use it and separately people with hearing impairments can use it, then those with both impairments may still be excluded.

#### **4.4 Survey structure**

Within the modules, it is important to consider the order of the items. If similar questions are asked in different modules, then they should be positioned consistently within those modules, to avoid undue influence from other items. For example, each module in the pilot contained a self-report question about limitations to daily activities. However, these were placed in slightly different locations in some of the modules. Placing them consistently at the start of each module would reduce variation in the influence of other items.

The length of the survey should also be considered. Some of the interviewers felt that the length of the pilot deterred some people from participating in it. However, this needs to be balanced with the need to include good-quality items on a wide range of capabilities. The module timings indicated that hand and arm function, vision and cognitive function were the longest modules (with mean completion times of 14 minutes, 9 minutes and 9 minutes respectively), leaving them with the largest scope for being cut down to reduce the length of the future full survey.

#### **4.5 Practicalities of the tests**

##### **4.5.1 Vision**

The post-processing revealed that several interviewers recorded the results of the vision tests incorrectly. This was mostly because of the complexity of the vision testing protocol where some participants read letters going up the chart and some read them going down the chart, depending on whether they read their initial row correctly. Interviewers were asked to record the largest row that participants got wrong. This was sometimes the last row the participants attempted and sometimes the row below this, depending on whether the participant was going down or up the chart.

The pilot survey thus indicated that the vision tests need to be simplified for use in the future survey. It is better if participants always start at the top of the chart and read downwards until they make errors. Clearer and more detailed computer instructions and interviewer training may also help. In particular, these should emphasize that a test is finished when the participant makes two errors on a line, not one.

Vision tests are more time-consuming compared to other performance tests. To keep the survey down to an appropriate length, it will be necessary to reduce the number of vision tests. An appropriate choice of tests would be a near vision high contrast test, conducted both with and without near vision aids, and a distance vision test with the best available setup of vision aids (if used). This will allow the closest comparison with standard vision data and act as a check and calibration of the vision results from the survey.

It is important in product use to measure some comfort vision levels as well as threshold ones as users are more willing to use products if they can use them comfortably. This should be done separately from the threshold tests to reduce confusion. It can be done by simply asking participants what is the smallest row that they find comfortable to read.

##### **4.5.2 Hearing**

The procedure used in the hearing tests could be shortened but only at the cost of increasing the complexity of the test and potentially confusing participants. It worked well in its current form and did not usually take that much time (most took 3-6 minutes, though one person took 17 minutes). Conducting the hearing test without hearing aids is justified as the hearing aid prevalence and usage is low, only 7.7% participants indicated they had a hearing aid, and of these, 5.7% of them said they rarely or never used it. It is recommended that the test could be used in its current form for the new survey.

##### **4.5.3 Hand and arm function**

Adjusting the grip strength dynamometer to fit the participant's hand was difficult and time consuming. More training for the interviewers to use the dynamometer may help to address the issue.

The dexterity product tasks seemed to go well with the equipment functioning as expected. Interestingly, the results highlighted a specific issue with the orientation of cables with participants getting confused about which way round the USB cables should be inserted. The mini USB cables used in the survey have now been effectively replaced with micro USB, but the issue of orientation remains and may even be worse for micro USB. A future survey could examine micro USB cables in order to determine what factors impact user performance.

#### 4.5.4 *Other issues*

Some of the items should be updated to allow better comparison with other surveys. For example, the options in the buying preferences module were based on (Philips 2004). However, they could be updated to match those in the new (Philips 2010) survey.

The interface style test was influenced by using a partial paper prototyping method. In this method, the interface is entirely static and does not respond to user input. This is likely to have influenced user behaviour. To get more accurate data, it would be better to use an interactive device. Touchscreen devices are now much cheaper than they were when the pilot survey was conducted, and could be programmed to prototype the interface styles. They could also be used for various other tests, such as vision, hearing, cognition and dexterity, to reduce the amount of equipment needed for the survey.

In addition, it may not be necessary to include both interface styles in the full survey. One interface style may be adequate as an indicator of participants' familiarity with technological interfaces.

## 5 **Conclusions**

This pilot has successfully delivered a proof of concept for product-related self-report questions and performance tests conducted by a national survey organisation in people's own homes. The data from this survey can be used to plan further nationally representative surveys and to pilot software and tools for informing product design.

An interesting finding from the survey relates to co-occurring ability losses. Of those participants who reported any limitations in daily activities due to capability loss, 44% of them reported limitations due to loss of more than one capability. This finding highlights the importance of measuring a range of capabilities in a single survey. The design, execution and analysis of the pilot survey presented in this paper provided valuable information on the various aspects of the survey. These include the sampling approach, survey structure, missing data, data distribution and feasibility of the tests, in addition to the appropriate levels of instructions, interview questions and response categories to be used in a future national capability survey. Some of the lessons learnt include (but are not limited to): adopting a stratified sampling approach; reducing the survey length especially in vision, hand and arm function and cognitive function modules; providing more training for interviewers; measuring multiple capabilities in a single survey; reducing exclusion criteria especially for Reach & Stretch tests; having back-up equipment; simplifying vision test protocol; starting hearing tasks at lower decibel levels; using a sole performance measure for grip strength; considering using micro USB for dexterity tasks and a single touchscreen device to measure multiple capabilities.

More importantly, what is gained from this work is the opportunity to engage the wider population directly in the design process instead of relying on product testers who may not necessarily reflect all levels of capability and interest. While it is not feasible to test all products with all groups consistently, incorporating this level of data – particularly in a full national study – into development ensures the opportunity for engineers to scientifically evaluate any possible barriers in their products. Only through these methods can more inclusive design be possible.

## **Acknowledgements**

This research was funded by the Engineering and Physical Sciences Research Council grant (Award Number: RG45089). Our sincere thanks to National Centre for Social Research for carrying out the field work under the

direction of the consortium. We also wish to thank Dr Jose Liht, John Ryan, and Timothy So for their analysis and literature reviews on the data collected.

## References

Blackler, A. (2006) *Intuitive interaction with complex artefacts*. PhD thesis, Queensland University of Technology, Australia.

Bradley, M., Waller, S., Goodman-Deane, J., Hosking, I., Tenneti, R., Langdon, P.M. and Clarkson, J. (2012) 'A Population Perspective on Mobile Phone Related Tasks' in *Designing Inclusive Systems (CWUAAT 2012)*, Springer-Verlag, London, pp. 55-64.

Bradley, M., Goodman-Deane, J., Waller, S., Tenneti, R., Langdon, P. and Clarkson, J. (2013) 'Age, Technology Prior Experience and Ease of Use: Who's Doing What' in *Ergonomics & Human Factors*, Cambridge, Institute of Ergonomics & Human Factors.

Clarkson, J., Coleman, R., Hosking, I. and Waller, S. (Eds), (2007) *Inclusive Design Toolkit*, Engineering Design Centre, University of Cambridge.

Clarkson, P.J. et al. (2012) *Towards Better Design, 2010* [computer file] Colchester, Essex: UK Data Archive [distributor], SN: 6997 , <http://dx.doi.org/10.5255/UKDA-SN-6997-1> (Accessed 1 June 2013).

Davis, A.C. (1995) *Hearing in adults. The prevalence and distribution of hearing impairment and reported disability impairment in the MRC Institute of Hearing Research's National Study of Hearing*. London: Whurr Publishers Ltd.

Dreschler, W.A. Verschuure, H. Ludvigsen, C. and Westermann, S. ( 2001) 'ICRA Noises: Artificial noise signals with speech-like spectral and temporal properties for hearing aid assessment', *Audiology*, 40, 148-157.

Elton, E., Johnson, D., Nicolle, C. and Clift, L. (2013) 'Supporting the development of inclusive products: The effects of everyday ambient illumination levels and contrast on older adults' near visual acuity', *Ergonomics*, 56:5, 803-817.

Erens, B., McManus, S., Field, J., Korovessis, C., Johnson, A.M. and Fenton, K.A. (2001) *National survey of sexual attitudes and lifestyles II*, Technical report. London: National Centre for Social Research.

Fennell, A.M., O Connor, J., Magennis, M., Fiddian, T., Bowden, C., Schmidt-Belz, B., Esser, A., Kirisci, P.T., Mohamad, Y., Gokmen, H. and Lawo, M (2011) 'Making Virtual Users a Reality: The Inclusive Design Project VICON', in *Include 2011*, Helen Hamlyn Centre, London.

Gershon, R. C., Cella, D., Fox, N. A., Havlik, R. J., Hendrie, H. C. and Wagster, M. V. (2010) 'Assessment of neurological and behavioural function: the NIH Toolbox', *Lancet Neurology*, 9(2): 138-139.

Goodman-Deane, J., Dong H., Langdon, P., and Clarkson, J. (2006) 'Increasing the Uptake of Inclusive Design in Industry', *Gerontechnology*, 5(3), 140-149.

Goodman-Deane, J., Waller, S., Williams, E., Langdon, P., and Clarkson, J. (2011) 'Estimating Exclusion: A Tool to Help Designers', in *Proceedings of Include 2011 conference*, Helen Hamlyn Centre, London.



Goodman-Deane, J., Waller, A., Collins, A.-C. and Clarkson, J. (2013). Simulating vision loss: What levels of impairment are actually represented? in *Ergonomics & Human Factors*, Cambridge, Institute of Ergonomics & Human Factors.

Grundy, E., Ahlburg, D., Ali, M., Breeze, E., Sloggett, A. (1999) *Research report 94 – Disability in Great Britain*, Corporate Document Services, London, ISBN 1-84123-119-3.

Hazel, C.A., and Elliot, D.B. (2002) ‘The dependency of LogMAR visual acuity measurements on chart design scoring rule’, *Journal of Optometry and Vision Science*, 79 (12): 788-792.

Huppert, F.A., Gardener, E. and McWilliams, B. (2006) ‘Cognitive function’, in Banks et al (Eds), *Retirement, health and relationships of older population in England: The 2004 English Longitudinal Studies of Ageing (Wave 2)*, Institute for Fiscal Studies, London.

*i~design Project*. [online] [www-edc.eng.cam.ac.uk/idesign3](http://www-edc.eng.cam.ac.uk/idesign3) (Accessed February 09).

Johnson, D., Clarkson, J. and Huppert, F.A. (2010) ‘Capability measurement for Inclusive Design’, *Journal of Engineering Design*, 21(2-3), 275-288.

Keates, S., and Clarkson, J. (2003) *Countering design exclusion: An introduction to inclusive design*, London: Springer.

Marmot, M., Banks, J., Blundell, R., Lessof, C. and Nazroo, J. (Eds) (2003) *Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing*, London: Institute for Fiscal Studies.

McGinley, C., Bound, J. and Dong, H. (2010) ‘Contextualising Data for Design’, in *Proceedings of the 5<sup>th</sup> Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT)*, Cambridge.

Molenbroek, J.F.M. (1987) ‘Anthropometry of elderly people in the Netherlands; research and Applications’, *Applied Ergonomics*, 18 (3), 187-199.

Ng, C.L, Ho, D.D. and Chow, S.P. (1999) ‘The Moberg pickup test: results of testing with a standard protocol’, *Journal of Hand Therapy*, 12: 309–12.

Norris, B. and Wilson, J. (1995) *CHILDATA: The handbook of child measurements and capabilities - Data for design safety*, London: Department of Trade and Industry.

Peebles, L. and Norris, B. (1998) *ADULTDATA: The handbook of adult anthropometric and strength measurements - data for design safety*. London: Department of Trade and Industry.

Philips (2004) *The Philips Index: Calibrating the Convergence of Healthcare, Lifestyle and Technology*, Philips Electronics North America, New York, NY.

Philips (2010) *Philips Index: America’s Health & Well-being Report 2010*, Philips Electronics North America, New York, NY.

Porter, J.M., Case, K., Marshall, R., Gyi, D.E. and Sims, R.E. (2004) ‘Beyond Jack and Jill: designing for individuals using HADRIAN’, *International Journal of Industrial Ergonomics*, 33, 249-264.

RNID ( 2013) *Check your hearing*. <http://www.actiononhearingloss.org.uk/your-hearing/look-after-your-hearing/check-your-hearing/take-the-check.aspx> (Accessed April 2013).

- Schwarzer, R. (1993) *Measurement of perceived self-efficacy. Psychometric scales for cross-cultural research*, Berlin: Freie Universität Berlin.
- Smith, S., Norris, B., and Peebles, L. (2000) *OLDER ADULTDATA: The handbook of measurements and capabilities of the older adult - Data for design safety*, London: Department of Trade and Industry.
- Snyder C. (2003) *Paper prototyping: the fast and easy way to design and refine user interfaces*, Morgan Kaufmann Publishers, San Francisco.
- Steel, N., Huppert, F. A., McWilliams, B. and Melzer, D. (2003) 'Physical and cognitive function', in Marmot, M. et al (Eds), *Health, Wealth and Lifestyles of the Older Population in England: The 2002 English Longitudinal Study of Ageing*, London: Institute for Fiscal Studies.
- Steenbekkers, L.P.A. and van Beijsterveldt, C.E.M. (Eds) (1998) *Design-relevant Characteristics of Ageing Users*, Delft University Press, Delft, Netherlands.
- Tenneti, R., Johnson, D., Goldenberg, L., Parker, R.A. and Huppert, F.A. (2012) 'Towards a capabilities database to inform inclusive design: Experimental investigation of effective survey-based predictors of human-product interaction', *Applied Ergonomics*, 43(4):713-726.
- The Welsh Health Survey (2003). <http://www.natcen.ac.uk/study/welsh-health-survey> (Accessed 20 April 2013).
- Van Dalen, D.B. (1979). *Understanding educational research* (4th ed.). New York: McGraw Hill, Inc.
- Waller, S., Goodman-Deane, J., Langdon, P., Johnson, D., and Clarkson, J. (2009a) 'Developing a Method for Assessing Product Inclusivity', in *International Conference on Engineering Design*, August 2009, Stanford, CA, USA.
- Waller, S., Clarkson, J., and Langdon, P. (2009b) 'Visualising design exclusion predicted by disability data: A mobile phone case study', in *UAHCI 2009, part of HCI International 2009*, San Diego, USA.
- Waller, S., Langdon, P., and Clarkson, J. (2010a) 'Using disability data to estimate design exclusion', *Universal Access in the Information Society*, 9(3):195-207.
- Waller, S., Langdon, P., and Clarkson P.J. (2010b) 'Designing a more inclusive world', *Journal of Integrated Care*, 18(4): 19-25.

**Table 1** Summary of capability measures, in %

<b>Vision</b>					
<b>Distance vision</b> (with vision aid if available) (Snellen visual acuity)	6/6 or better	6/7.5 – 6/12	6/15 or worse		
	69.3	26.8	3.8		
<b>Near vision</b> (high contrast letters, with vision aid if available) (size of smallest letters read)	0.6–0.8mm	1.0-1.2mm	1.45-1.8mm	2.3mm or larger	
	60.7	27.7	6.6	5.0	
<b>Hearing</b>					
<b>Hearing at different volumes</b> (with no background noise) (lowest volume heard correctly)	Low volume (25dB)	Medium volume (40dB)	High volume (60dB)	None of the volumes	
	83.0	13.2	3.2	0.6	
<b>Hearing at medium volume</b> (with different levels of background noise) (highest level of background noise heard correctly)	High noise (60dB)	Medium noise (40dB)	Low noise (25 dB)	No noise	
	42.8	52.3	2.9	2.0	
<b>Hand and arm function</b>					
<b>Moberg</b> (time to pick up 12 small objects)	0-14sec	15-19sec	20-29sec	30sec or slower	
	65.7	25.6	5.6	3.1	
<b>Threshold grip strength</b> (dominant hand) (men)*	49kg or greater	43-48.9kg	34-42.9kg	Less than 34kg	
	25.3	26.5	24.7	23.5	
<b>Threshold grip strength</b> (dominant hand) (women)*	30kg or greater	26-29.9kg	21-25.9kg	Less than 21kg	
	24.2	24.7	29.2	21.9	
<b>Comfort grip strength</b> (dominant hand) (men) *	41kg or greater	31-40.9kg	22-30.9kg	Less than 22kg	
	24.1	27.2	23.5	25.3	
<b>Comfort grip strength</b> (dominant hand) (women) *	25kg or greater	20-24.9kg	15-19.9kg	Less than 15kg	
	25.3	24.7	24.7	25.3	
<b>Successful completion of everyday dexterity tasks</b>	Pressing small buttons	Pressing medium buttons	Inserting headphone jack	Inserting USB cable	
	77.9	95.2	87.4	99.2	
<b>Reach and stretch</b>	Able to fully extend both arms in front and above head	Able to fully extend one arm but not the other	Unable to fully extend either arm	Not asked due to a recent back or shoulder injury or operation	
	71.8	2.8	2.2	15.7	
<b>Mobility</b>					
<b>Getting out of a chair with arms</b>	Never difficult	Occasionally difficult	Frequently difficult	Always difficult	No experience of this task
	80.4	12.4	4.4	1.4	1.4
<b>Getting out of a chair without arms</b>	76.5	14.6	3.0	4.7	1.1
<b>Reaching floor level</b>	74.0	16.3	4.1	4.1	1.4
<b>Cognitive function</b>					
<b>Memory recall (immediate plus delayed) (words recalled out of 20)</b>	14 + words	12-13 words	9-11 words	0-8 words	
	27.0	18.8	30.1	24.7	
<b>Visual search task: Search speed</b>	400 + letters	300-399 letters	250-299 letters	0-249 letters	
	17.6	31.2	27.2	24.0	
<b>Visual search task: Search efficiency</b>	95% or higher	85%-94%	70-84%	Lower than 70%	
	16.5	39.3	29.2	15.0	

<b>Literacy (three questions)</b>	All answers correct	2 out of 3 correct	1 out of 3 correct	None correct	
	59.8	24.2	12.8	3.1	
<b>Numeracy (four questions)</b>	All answers correct	3 out of 4 correct	2 out of 4 correct	1 out of 4 correct	None correct
	55.7	25.2	14.0	3.9	1.1

\* Results for men and women are %s of the male and female sample only.

**Table 2** Limitations in daily activities due to capabilities, in%

	Not at all limited	Somewhat limited	Very limited
<b>Vision</b>	89.5	9.4	1.1
<b>Hearing</b>	94.5	5.0	0.6
<b>Hand/Arm*</b>	82.9	14.9	2.2
<b>Mobility</b>	84.5	13.3	2.2
<b>Cognition</b>	84.8	14.1	1.1

\*Participants were categorised as having a hand/arm limitation if they reported a limitation in either or both dexterity and reach & stretch.

**Table 3** Use of technology products, in %

Technology task	% of sample does this task Frequently or Occasionally (as opposed to Never or Rarely)	% of those who had done the task and found it Difficult, Very difficult or Impossible
Calling on a mobile phone	77.5% (n=360)	3.0 (n=332)
Texting on a mobile phone	67.8% (n=360)	6.0 (n=284)
Taking pictures with a digital camera or mobile phone	59.4% (n=360)	3.5 (n=285)
Using a remote control for digital TV	93.6% (n=359)	2.9 (n=347)
Using the internet	73.6% (n=360)	3.5 (n=283)
Listening to MP3 tracks on a portable device	35.8% (n=360)	4.0 (n=175)
Using gaming console	21.1% (n=360)	7.1 (n=140)
Using satellite navigation	28.6% (n=360)	10.4 (n=154)

**Table 4** Most important consideration in buying a new technology product, in%

Quality	Ease of use	Range of functions	Low price	Appearance	Durability	Environmental impact
31.8	22.5	17.5	16.1	5.6	5.1	1.4

**Table 5** Self-reported height and weight

	Men	Women
Mean height (cm)	176.2	164.0
Mean weight (kg)	82.5	68.6

**Table 6** Co-occurrence of self-reported limitations, in %

<b>Condition (N=362)</b>	<b>% of sample</b>
<b>No Limitations</b>	<b>62.7</b>
<b>Any Limitations</b>	<b>36.3</b>
Vision	10.5
Hearing	5.5
Mobility	15.5
Cognitive function	15.2
Hand and arm function	17.1
<b>One limitation only</b>	<b>21.0</b>
Vision only	4.4
Hearing only	2.2
Mobility only	3.6
Cognitive function only	6.1
Hand and arm function only	4.7
<b>Two or more limitations</b>	<b>16.3</b>
Two limitations	8.6
Three or more limitations	7.7