Interacting with Computers 22 (2010) 510–529

Contents lists available at ScienceDirect

Interacting with Computers

journal homepage: www.elsevier.com/locate/intcom

Developing a model of cognitive interaction for analytical inclusive design evaluation

Patrick Langdon [a,*], Umesh Persad [b], P. John Clarkson [a]

[a] Engineering Design Centre, Cambridge University Engineering Department, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom
[b] Centre for Production Systems, University of Trinidad and Tobago, Trinidad and Tobago

ARTICLE INFO

Article history:
Available online 16 September 2010

Keywords:
Inclusive design
Universal access
Cognitive psychology
Cognitive modelling
User testing
User modelling

ABSTRACT

Inclusive design is a user-centred approach that examines designed product features with particular attention to the functional demands they make on the perceptual, thinking and physical capabilities of diverse users, including those with impairments and ageing. An analytic approach to the evaluation of designs mitigates the need for observational trials with products by relating data about the prevalence of capability ranges in the population with an analysis of the demands made by product properties and features. This enables a quantification of the number of users who can use a specific design. To date, there has been some success in identifying data sets and appropriate impairment and capability models for perception and movement in this novel "inclusive" research context. However, previous attempts to do so for cognitive aspects of product feature interaction have encountered a lack of suitable data and models. We propose some necessary requirements for a complete model of inclusive cognitive interaction and establish four criteria for what would constitute a good framework for the purpose of developing a research approach that could be used to construct and test predictive tools for design. Taking into account the immediately relevant literature, we examine some candidate approaches that may satisfy these requirements with reference to some of our own research findings. The results of the analysis suggest that this combined approach to cognitive demand is, in principle, capable of satisfying the proposed criteria in conjunction. It has also been successful in driving a research effort to identify important predictive variables and relate these to an underlying model of interaction. The utility of such a framework will ultimately be judged by empirical tests of the accuracy of the developed model and tools in predicting specific exclusion and difficulty during cognitive interaction. This will allow further iterative improvement of the model and will also permit modification of the development framework. A further test will be whether designers can use the resulting tools to help create designs for ICT products that are more inclusive in that they are usable by people with a wider range of functional capabilities.

1. Introduction: analytical inclusive design

The relationship between the user's capabilities and their environment has long been considered in the context of assistive technology. Newell (1995), for example, considers the equivalence of ordinary interactions impaired by extreme situations and impaired individuals in ordinary situations. Design should focus on the extra-ordinary or impaired first, accommodating mainstream design in the process (Newell and Gregor, 1997).

Many products today are laden with a host of features which for the majority of users remain unused and often obscure the use of the simple features of use for which the product was devised. For example, since the target cognitive capabilities anticipated by the designers are often similar to their own demographic and largely not affected by age-related cognitive impairment, the cognitive demands made by such products are frequently high.

Inclusive design aims to make products and services accessible to the widest range of users possible irrespective of impairment, age or capability. To do this, a substantial research effort has been directed towards developing the underlying theory and practice of design analysis in order to develop and provide tools and guidance to designers that they can use to improve the inclusion of a resulting product (Coleman, 2001; Clarkson et al., 2003; Nicolle and Abascal, 2001). This research has principally been motivated by changes in national and International demographics evident in recent years leading to increasing numbers of older and reduced numbers of younger product users. The problems that arise from this include increased prevalence of impairment, which has led to legislative requirements for greater inclusion at home, at leisure and in the workplace.

* Corresponding author. Tel.: +44 1223 766959; Mobile: +44 7767 880815; fax: +44 1223 332662.
E-mail addresses: pml24@eng.cam.ac.uk, pjc10@eng.cam.ac.uk (P. Langdon), umesh.persad@utt.edu.tt (U. Persad).

0953-5438/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2010.08.008

Inclusive design, as developed by this approach, examines designed product features with particular attention to the functional demands they make on the perceptual, thinking and physical capabilities of diverse users (Vanderheiden and Vanderheiden, 1992; Freudenthal, 1999; Clarkson et al., 2003; Keates and Clarkson, 2004; Persad et al., 2007). Other related approaches, for example, use a single comprehensive data set from 100 people containing anthropometric and behavioural data but omit cognitive information. This latter HADRIAN system allows designers to query the data base for numerical and personal information (Porter et al., 2004a) and the SAMMIE development allows the data to be used in a task analysis with trials of fit between virtual human figures and interfaces (Porter et al., 2004b).

This sort of analytic approach to the evaluation of designs is intended to mitigate the need for extensive observational trials or experiments with products in use by selected samples of people. Instead, data about the prevalence of capability is used holistically in conjunction with a categorisation and analysis of product properties and features, in order to quantify the numbers of users and sub-populations of users, who can use a design without mental or physical failure or difficulty. Such an inclusion "audit" of a design requires: (1) a data set of the distribution of capability in the population, and (2) methods for characterising and measuring the products' demands in different functional areas; sensory, cognitive and physical (Persad et al., 2007). However, it is also clear that it must take into account situational factors such as the social, physical and economic context of use.

Importantly, this latter approach is a novel research development related to design for all and universal design (Story et al., 1998), that focuses on the exclusion of individuals from using products as a result of difficulties of use or failure of interaction resulting from the product's demands exceeding users' capability. The quantification of this capability–demand relationship is therefore referred to as a calculation of product inclusion, and the application of its principles to cognitive interaction, for example, is here referred to as "assessing the degree of cognitive inclusion" (Fig. 1).

2. The quantitative approach to inclusion

Initial research focussed on the only known complete set of data on capability variation for the population publically available through the UK Office of National Statistics (Keates and Clarkson, 2004). This representative national data was from a UK disability follow-up (DFS) survey of over 7000 individuals carried out in 1997 and was intended to establish the prevalence and severity of quality of life problems arising from functional impairments (DFS-SN4090, 1997). It used a methodology devised for a number of partial surveys of Family resources and disability carried out by the UK Office of Population Censuses and Surveys (Grundy et al., 1999; Martin and Elliot, 1992). The 1996–1997 survey used a self-report scaling approach, where respondents answered question items that were intended to locate their levels of functional impairment on a set of scales that were then used to calculate an overall index of severity of disability. The scales used included specific question items addressing functional capabilities, such as vision, hearing, dexterity, reaching and stretching, locomotion, intellectual function and communication. A number of scales related to areas such as continence, digestion and scarring were less relevant. Despite a mismatch between these disability based scales and the requirements of functional capability analysis, recent research developments have successfully deconstructed this data set enabling derived values to be used in a psychological scaling approach to a product audit procedure. Hence, visual demand can be ascertained by asking designers to estimate the visual demand of the product feature on a visual scale with anchor points describing viewing text of different sizes (Waller et al., 2010a). Similarly, physical demand may be estimated by making comparison judgements with question items involving picking up pins or lifting objects and walking set distances. These judgements are indexed to the survey items of the original data set to yield estimates of numbers of individuals excluded at a given scale level (Waller et al., 2010b).

2.1. Quantifying cognitive capability

The DFS survey scale items were compiled from sets of items devised by a chosen set of knowledgeable judges. One weakness of this approach, however, was the poor match of the Intellectual function survey items with cognitive capability. In particular, the survey scales were defined by judges who were practitioners and therapists, and focused on complex everyday tasks requiring intellect, rather than on current theories of cognitive function (Martin and Elliot, 1992). Furthermore, the resulting lack of correlation between items and scales and between individual judges led the survey designers to construct a scale that simply summed the

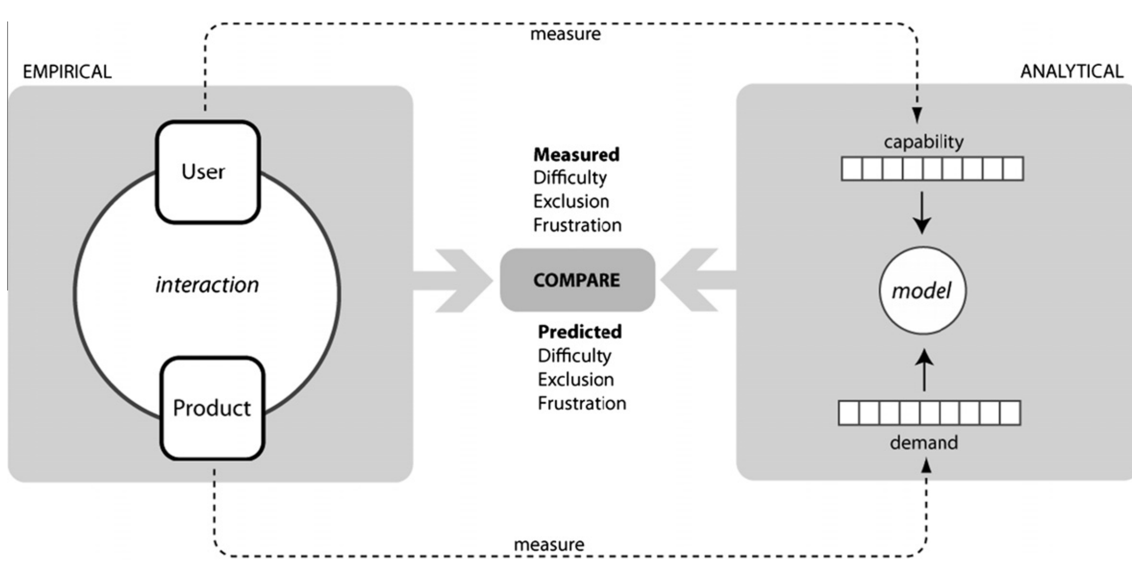

Fig. 1. The analytical approach to evaluation and empirical validation.

numbers of intellectual problems selected form a undifferentiated list. An analysis of the requirements for analytical inclusive design scales carried out by Persad et al. highlighted this weaknesses of the survey scale functional assessment approach and discussed the requirements for improved capability measures (Persad et al., 2007). For example, the survey's visual scale gave estimates of the individual's global level of visual functioning along a visual capability dimension. This was done by combining the responses to graded sets of items that addressed particular functions selected by practitioners, such as reading and person recognition. Persad et al. (2007) hypothesised that a more accurate approach would be to use objective measures of low-level functional capability, in conjunction with a model relating these low-level capabilities and their combination. This would be used to predict high-level task performances, such as recognising a dial setting, pressing a button or selecting a menu item.

An alternative approach would be to use psychological scaling as a method of obtaining precise estimates of visual and motor functional demand from specific product features. This has been investigated in more recent development of the inclusive design scales in order to quantify sensory and physical functions (Waller et al., 2010b). However, the research described here takes a different starting point: namely, that of developing a predictive model in the cognitive functional domain. This approach was outlined by Persad et al. (2007), and we describe new developments of this theory and methods in the light of associated empirical findings.

2.2. Ageing and cognitive capabilities

Rather than consider impairment as synonymous with a disability or medical condition, the casting of cognitive impairment as a set of interrelated functional capabilities is an approach that has been introduced in the context of designing information technology for cognitive support. The idea here is that cognitive dysfunction due to ageing and other causes can be reconceptualised, as part of an inclusive user-centred design process, into various qualitative types of capability and the ranges for them that can be found in the population (Newell et al., 2008). Newell et al. utilised a number of general categories from information processing models in cognitive science, including: Intelligence, perception, memory, attention and response generation. They identified specific issues of relevance to the design of ICT interfaces, including sensory decline in terms of speed; the avoidance of complexity; simplification of language and the streamlining of interaction tasks.

It is known that cognitive capabilities such as verbal and visuospatial IQ show gradually decreasing performance with aging (Schaie, 1988; Rabbitt, 1993; Petrie, 2001; Nichols et al., 2006; Newell et al. , 2008). Attending to goal-relevant, task features and inhibiting irrelevant ones is important in product interaction and this is known to be affected by ageing. Attentional resources may also be reduced by ageing, such that more mistakes are made during divided attention, dual task situations. This will affect the individual's ability to deal with the cognitive workload demanded by the use of a product, particularly when individual cognitive performance in terms of memory or speed of processing can lie within a wide range.

Specific conditions such as stroke, dementia, autism, Asperger's syndrome and cerebral palsy may differentially affect normal population ranges of capability. This can further reduce visual–spatial perceptual capability and memory, as well as impair the comprehension of text and speech in aphasias. Language capability may additionally be affected by variations resulting from speech impairments, literacy levels and dyslexia. Finally, co-occurrences of perceptual, cognitive and physical impairment introduce further complexity and variability, suggesting that design for cognitive impairment should be strongly sensitive to individual needs (Newell et al. 2008). Freudenthal (1998) identified specific problems that older users may have had as a result of lacking the skills and abilities required for modern product functions. Other deficits identified included parallel processing capability and attentional resources, suggesting that products should be specifically designed to reflect the older users' experience generation. In the absence of suitable declarative knowledge in long-term memory (LTM) to enable reasoning about product features, users may decline to interact, or resort to trial-and-error manipulation of product features. This amounts to trying actions on random interface elements, looking for those that produce results that match to previously encountered procedural schema (Blackler, 2006; Langdon et al., 2007, 2010).

These considerations suggest that in order to develop a model of product cognitive demand to adequately address the range of cognitive capabilities in the population it is important that a complete range of capability variation and its causes should be considered and employed in developing and testing a research framework.

3. Quantifying cognitive demand

The overall aim was to create a framework suitable for the purpose of developing a research approach to address the quantification of cognitive demand and capability measurement. This quantification was required in order to develop and validate methods that will be used to underpin predictive techniques. These could in turn be rapidly employed in industrial settings to indicate aspects of designs that may create cognitive demand that exceeds the capability of individuals in the desired user population. Fig. 2 outlines the strategy adopted for development of such a framework. The principal requirement was to quantify the relationship between product demand and user capabilities. This need was then used to identify key theoretical barriers, problems and issues that would have to be addressed for success and also to specify a candidate set of requirements for an adequate framework for developing a model of cognitive demand. The theoretical issues, such as the need to accommodate the effects of prior experience, were also used to assess the applicability of candidate models of interaction. The requirements for an adequate framework were then used to define the key elements of the model and the operational variables that would need to be operationalised and measured in order to predict performance. The specific theoretical problems and chosen model gave rise to a derived research strategy, while the considerations of the model and operational variables provided the actual research questions and experimental designs. This framework approach follows that proposed by Jacko and Vitense (2001) in the wider context of user profile modelling to assist the design of ICT for disability.

Finally, the outcomes of the research were expected to yield models of testable methods for predicting exclusion and difficulty from demand and capability. The qualitative and quantitative shortfall found in these of these could then be established and the insights gained used to modify the initial theory and requirements.

3.1. The requirements criteria and framework

As a starting point, we examine a systematic set of requirements criteria for what would constitute an adequate framework for research into inclusive cognitive demand. We propose the following:

- An accepted underlying model of cognition that encompasses the perception, cognition and action loop of interaction (Fig. 3);
- A set of identified empirical variables and measurements that can be used to establish the effect of key cognitive functions such as memory and attention, in inclusive interaction;

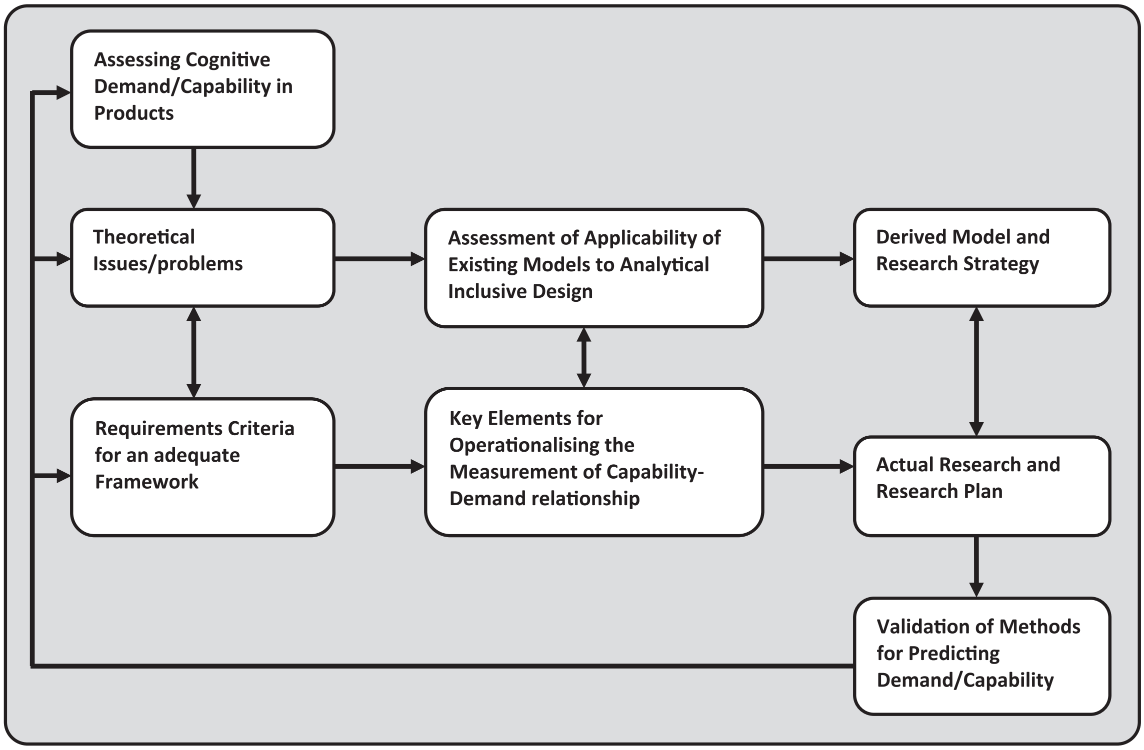

Fig. 2. Development of a framework for cognitive analytical inclusive design research.

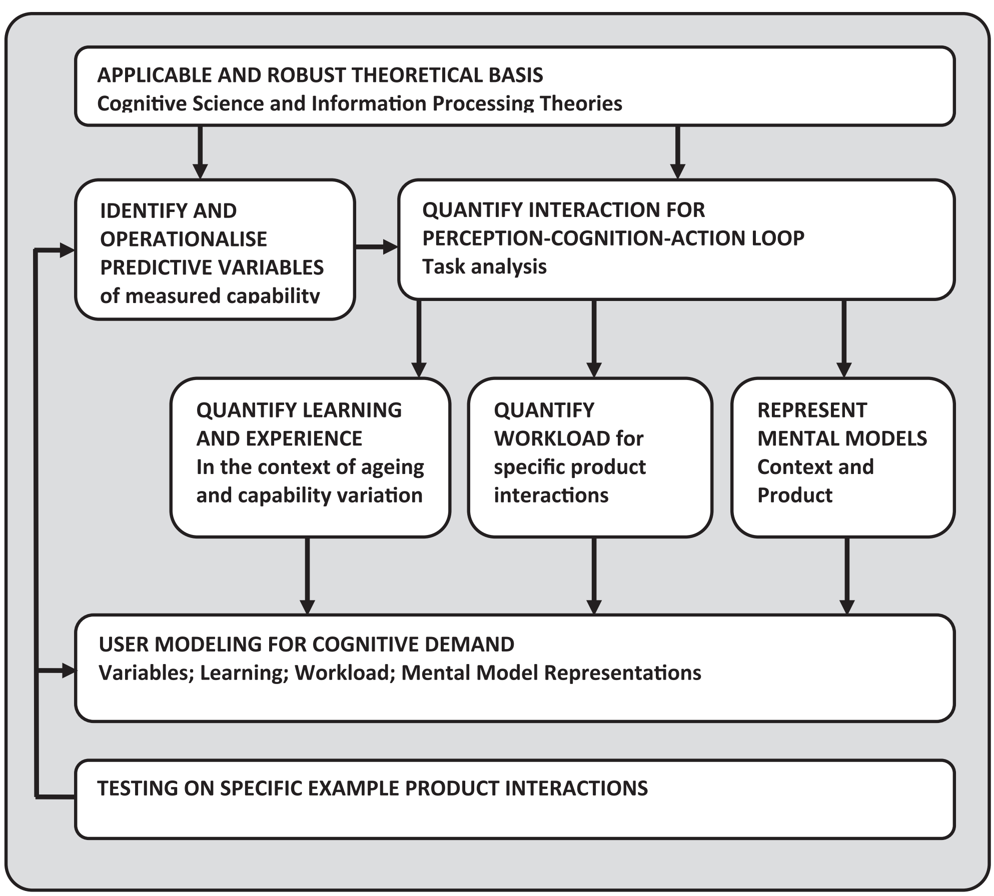

Fig. 3. Components of a framework for analytical inclusive design research.

- A demand measurement method that can accommodate the extent of prior experience in product users and the transition from novice to expert user;
- A user modelling system that can instantiate the cognitive theory and operationalise the inclusive, demand–capability relationship.

Any approach to cognitive demand that satisfied these criteria in conjunction could be utilised to identify important predictive variables and would be able to relate these to an underlying model of interaction. The power of the model and associated variables for predicting inclusive performance could be tested by comparing real user interaction data with the model's predictions. Capability data could then be collected from the population at large that corresponded to the identified predictive variables and the quantitative model's predictions could be matched to new representative population data sets giving the prevalence of cognitive capabilities. The combination of representative samples of specific quantitative data from the UK population and advanced interaction modelling would constitute a powerful tool for calculating cognitive demand from product features.

We examine how alternative theoretical approaches can accommodate and predict key aspects of product interaction and attempt to delineate the areas where they cannot and where the effective modelling boundaries occur. To do this we will address:

- different information processing models of essential cognitive function;
- the measurement of multimodal workload in multiple tasks;
- human factors models of the acquisition of skills for high performance tasks;
- the criteria for representations of users' and designers' mental models for inclusive design.

Some data and conclusions from research into the relationship between various cognitive variables and user's performance with products will then be examined for the light it can shed on methods for measuring inclusion demand. The overall aim will be to establish the elements that are necessary to create the empirical and theoretical groundwork for an approach to quantifying inclusive cognitive interaction. In the following sections, we will address the issues behind each criterion in turn with reference to some relevant empirical findings that have been generated during the progress of the developing research strategy.

4. Criterion 1: an accepted underlying model of cognition in interaction

The requirement is for a generally accepted and robust base theory, derived from experimental findings in cognitive psychology. It should be capable of identifying significant variables related to capability and provide a rationale for operationalising these into measurable quantities that can be used to predict interaction performance. It should also provide the theoretical elements usable in a model to quantify interaction. These elements will include memory, attention and reasoning functions concerned with the processing of perceptual information and the resulting initiation of actions.

4.1. The disability survey model

The disability survey data approach simply introduced a number of items into the survey that described typical problematic cognitive difficulties. These were devised by a panel of judges and were based on quality of life metrics suggested by practitioners (Martin and Elliot, 1992). The number of positive responses to these unordered items were simply summed and rescaled to produce an 'intellectual function' score:

1. Hold a conversation without losing track of what is being said.
2. Do something without forgetting what the task was whilst in the middle of it.
3. Think clearly, without muddling thoughts.
4. Count well enough to handle money.
5. Tell the time of day, without any confusion.
6. Watch a 30 min TV program, and tell someone what it was about.
7. Read a short newspaper article.
8. Write a short letter to someone without help.
9. Remember a message and pass it on correctly.
10. Remember to turn things off, such as fires, cookers or taps.
11. Remember the names of friends and family that are seen regularly.

A separate communication scale covered the expression and understanding of speech (Grundy et al., 1999). It seems likely that this assessment scale does not reflect cognitive capabilities in a systematic way that relate clearly to cognitive theory. It is dependent on diverse original test items that confounded a range of cognitive functions, In particular, the items do not map directly to product interfaces and often confound a number of underlying cognitive functions such as memory, visual processing and attention. There are no clear mappings between low-level functions and high-level tasks due to the vagueness of item descriptions.

4.2. An applicable structure from cognitive science

It was important that the chosen theoretical basis be clear and simple but powerful for modelling. It should be able to generate, in a transparent way, the key elements of quantitative demand along with associated operationalised research variables (Fig. 4). The key elements were identified as:

- A quantitative information processing model of the interaction loop consisting of: input processing; short and long-term memory; executive function for decision making; and output responses;
- The ability to underpin quantification of workload in the form of varying cognitive demand made by serial and parallel task requirements;
- It should enable quantitative modelling of the effects of prior experience and learning on interaction with products;
- It should aid the specification of mental models to help quantify interaction with defined product features.

Considerations of cognitive information processing during sensory input would need to be underpinned by comprehensive theories such as that of Marr's (1982) account of early visual processing. A overview of higher cognitive processes based on the human information processing approach would include functions such as attention; sensory memory; short term memory (STM); working memory (WM); long-term memory (LTM), as well as memory models implying organising processes, such as semantic or declarative, episodic, and procedural memory. Following current theories, such as that of Baddeley (2000) or Wickens (2002) it would also assume an executive function that manages the operations of working memory. Such models generally accommodate the requirements of instantaneous attention and its interaction with material stored more permanently for a longer term (Fig. 4). This model is central to the idea of situational or contextual cues

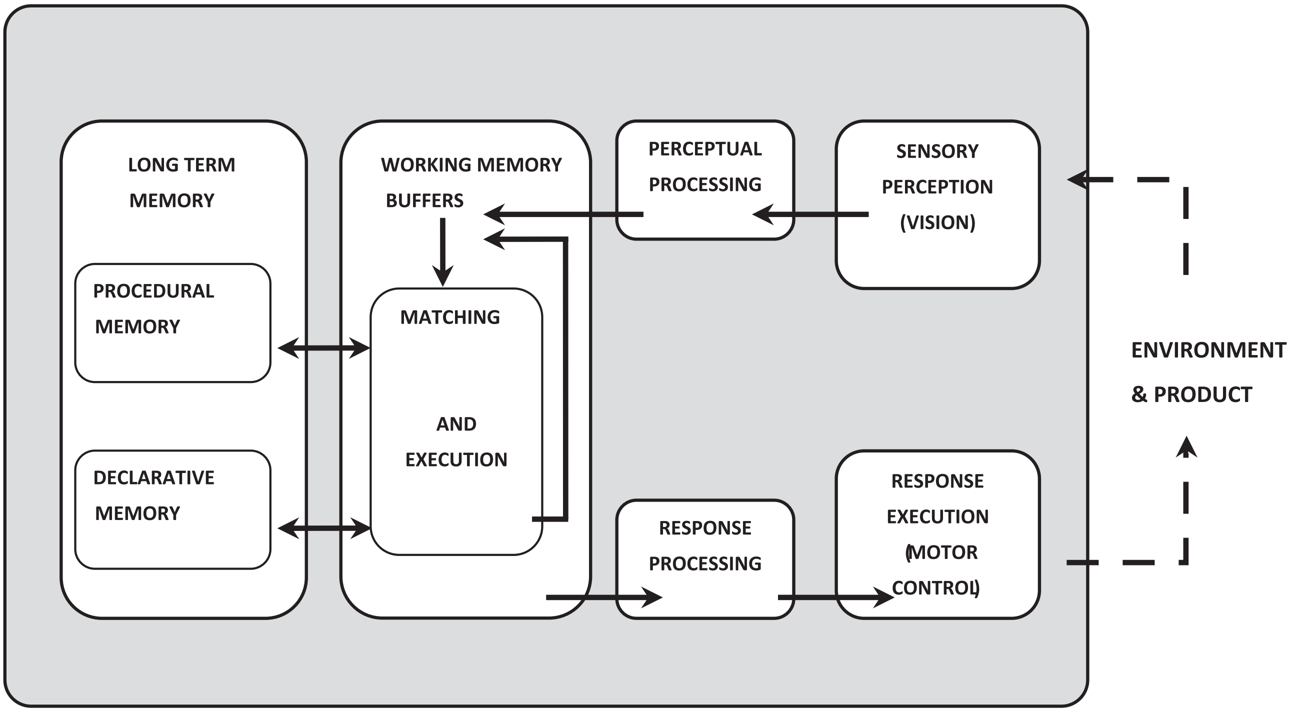

Fig. 4. A candidate cognitive model for inclusive interaction based on widely accepted cognitive science.

being processed in WM, matching to and retrieving schemata or skilled responses from LTM in order to initiate actions. It is assumed that the most similar and frequently occurring matching schemata are activated first (Anderson, 1982). The collection of available LTM memory structures that are available to match to processed cues from the product interaction are assumed to form a body of unconscious predispositions to act in a specific way in given circumstances. They are further assumed to be stored in context specific form. A structure of this sort has been proposed as a framework for research into human error and is widely accepted in cognitive science (Reason, 1990; Anderson, 2007).

4.3. Experience and learning models

Experience is a critical factor of how easy a product is to learn. All products bear some resemblance to products extant during previous generations or products from different companies or product families. It follows that new products that use this resemblance to help the user make a reference to the same function on a familiar device should outperform those that make no such association. Memories relating to the experience of products will be stored in the long term store and the ability of the central executive to find the relevant knowledge will depend on the cues available and the level of previous experience. In general, the more times a user has completed a task the faster this process will become as they progresses from novice to expert.

The effect of prior experience is known to be most negative when the same interface features are associated with different responses. On the other hand, a variety of differently appearing product interface features does not affect responses when the basic layout of controls is the same. Interestingly, the degree of similarity of a product that has been experienced before does not necessarily have to be high to allow effective learning (Wickens and Hollands, 2000). There is evidence that the existence of particular perceptual consistencies and the presence of key task elements alone may be as good as realistic simulations for pilots (Lintern et al., 1990). Guided training research has also shown that reducing the number of available menus in a word-processor in the early stages of learning, accelerates learning and reduces the impact of errors. This suggests that reducing the number of errors that users experience during early prior experience of a product facilitates the learning of a task with it, although errors are necessary for learning to take place (Carroll and Carrithers, 1984). It has also been argued that tasks with similar demands generate expectations of similar responses or actions. Such a "task-action consistency" suggests that product interfaces and usage situations that do not take into account the effects of prior experience will be unusable and also difficult to learn (Ryu and Monk, 2004).

4.3.1. Skill acquisition

It is assumed that during the transformation from novice to skilled performance, an individual moves from predominantly declarative, knowledge-based performance to more procedural, rule-based and skill-based performance and that this is accompanied by reduced cognitive workload. This is well accounted for by models based on theories of human associative memory such as ACT (Anderson, 2007). It is also an established experimental finding. Rasmussen, in studies of electrical fault finding, (1986) distinguished three levels of processing in the selection of actions:

- *Skill-based*: A rapid automated response to key feature cues;
- *Rule-based*: The retrieval of action schemas into WM in order for some slower 'if–then' reasoning to take place;
- *Knowledge-based*: Unique goal-based reasoning about the situation and perceived symbolic cues, based on declarative mental models and general purpose propositional reasoning rules.

Rasmussen's model has since been cited as a major antecedent of Human Action Theory and the Action Facilitation approach (Arnold and Mettau, 2006). In these accounts, desirable outputs of interactions between the user and the artefact within a given environmental context are a function of the user characteristics, the task characteristics, and the properties of the artefact. A good match of these elements in terms of cognitive, perceptual and motor processes is associated with effective, efficient and desirable outcomes. This family of models are related to the framework being developed. In particular, they also describe the selection and execution of an action programme or routine at the skill, rule or knowledge-based levels and give an account of how this execution is modified by feedback of results. The models propose that the design of artefacts should support the optimisation of performance in a work setting by minimising learning and maximising affordances for use.

Similar representational levels have been invoked to account for so-called intuitive interaction. For example, Norman (1986) developed a theory of how knowledge in the form of mental models may underlie individual interactions with everyday interfaces, particularly when the interactions are faulty as a result of poor design (Norman, 2002). More recent investigations into the effects of older individuals' prior experience during product interaction have found evidence for a more significant role for skill and rule driven interactions than for knowledge-based exploration (Blackler et al., 2003; Langdon et al. 2007, 2010).

4.3.2. Generational effects

For older generations some modern symbols used in ICT products can go completely unnoticed (Lewis and Clarkson, 2005). Studies in the Netherlands have explored this through the concept of technological generations, characterised by the types of interface individuals were familiar with when young. Docampo-Rama (2001) showed that this affects the learning of interfaces such that performance errors and task timings may show discontinuities between differently aged generations in their interaction with modern technology products. In particular, layered computer interfaces have been shown to be best suited to the learnt processes of those aged 25 and younger. This is particular relevant to some of today's consumer products that have features of use that are unfamiliar and poorly matched to older users' experience. There may also be more recent, technology generations whose experience is shaped by post-mouse input devices, such as game pads and touch screens and sophisticated error-tolerant user interfaces such as those found in computer games, PDA's and mobile phones. New developments in computer interaction, including avatars, ubiquitous and ambient computing and multimodal voice and gestural interfaces, would be expected to form the foundations of further generational effects in the future population. The extent and nature of users' prior experience as well as their membership of possible technology generations will need to be accommodated by any predictive model for analytical inclusive design.

4.4. Mental models

We propose that a representation of the users' mental model of a device could assist the estimation of cognitive demand by providing the declarative and procedural information for a task breakdown necessary for cognitive workload analysis. Mental models are generally taken to be internalised mental representations that reflect and can be used to reason about interaction in the external world (Rogers et al., 1992). They may be coherent analogues of the real world or may consist in incorrect or incoherent fragmented beliefs, for example about the mechanical or electrical physical behaviour of objects (Gentner and Stevens, 1983). Mental models are likely to capture the knowledge required in planning, evaluating and executing actions during interaction. However, they may be incomplete, inaccessible to intuition, incorrect or generic to a number of alternative situations (Norman, 1986). Past research has characterised mental models as capturing the semantics of thought (Johnson-Laird, 1980) used for deductive reasoning. Alternatively, they may represent syntactic information in the form of rules or procedures to be activated in given circumstances in order to predict events (Kieras and Bovair, 1984; Payne, 1992), particularly the interaction with artefacts such as designed products or computer interfaces. More recent findings reviewed by Payne (2008) suggest that mental models may also consist of an elaborated problem space within which reasoning can also occur during interaction. This is assumed to be consistent with the knowledge-based level of processing of interaction as proposed in Rasmussen's (1986) goal-directed, action hierarchy.

We assume that mental models are based on a particular user's previous experience. It is also assumed that they are continuously modified through interaction and that they are generated in working memory using information stored in long-term memory matched to perceived task, environmental and artefact characteristics. In particular, the approach taken assumes that interaction involves a dialogue between the changing behaviour of the product and the objects and actions stimulated in the user's mental model representation as a result (O'Malley and Draper, 1992).

The requirements for a representation that captures the role of mental models in cognitive interaction are apparently straightforward. For example, Fig. 5 shows a state space representation for the reactive behaviour of a toaster using a state chart based on the unified modelling language. State charts are also commonly used to specify the design of reactive systems and are familiar to analysts and designers of embedded systems (Thimbleby, 2007).

The demanded user action sequence for making toast is also shown in Fig. 6, consisting of six steps. 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. Therefore, one approach to the problem of assessing cognitive workload is to attempt to construct the mental model demanded by a specific task. Using this model as a representation the aim is to evaluate the level of declarative knowledge required to perform the task, and the problem solving, planning and search that is required to construct a path through the state space of the device and user as they interact.

The representation can be used to derive the action sequences that form users' mental models by analysing variations to the demanded action sequence.

Because working memory is limited, it follows that there will be a limit to the complexity of the mental model and the mental operations that can be performed with it. Therefore, for users with lower cognitive capability, there will be a limit to the complexity of the mental model that could be used effectively during interaction and this will give rise to difficulty in using the product and hence the exclusion of some individuals from the use of it.

4.5. Workload and cognitive resource theory

While physical factors have long been quantified in classical ergonomics, human mental workload has proved a more difficult phenomenon to derive. It has been quantified using approaches such as the NASA-TLX task-load index (Hart and Staveland, 1988; Arnold, 1999). However, this is dependent on subjective ratings and tested on performance tasks, rather than operationalising an underlying model of cognition in interaction. To account for human attentional performance resulting from limitations of cognitive resources in overload situations, Wicken's multiple resource theory of workload has been developed specifically to account for high performance interaction environments such as air-traffic control or flying fast jets (Wickens, 2002). The theory predicts performance using quantification of the limitations and distribution of cognitive resources in multi-task interactions. It models cognitive resources as being dependent on four dimensions derived from findings in experimental psychology:

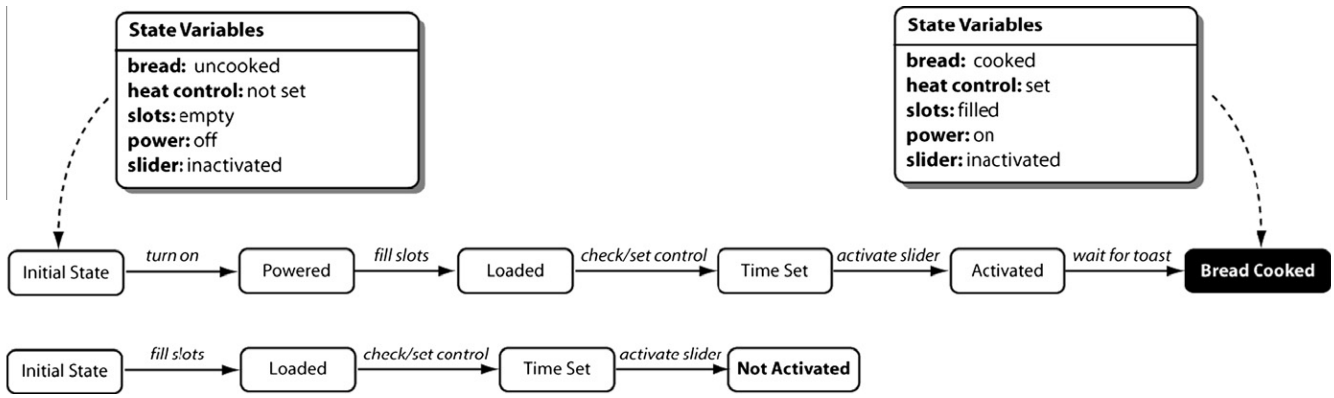

Fig. 5. A state space representation can capture incomplete and incorrect mental models.

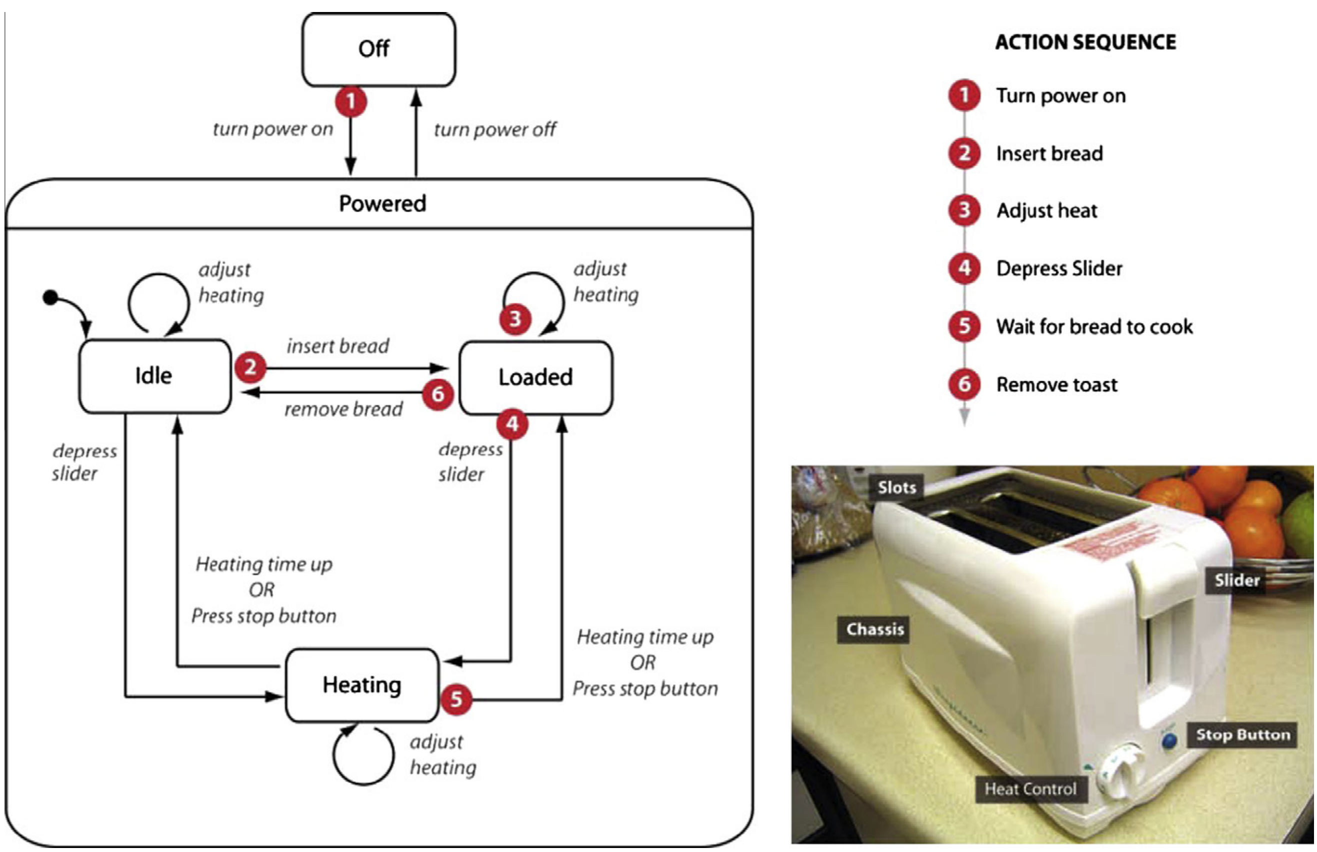

Fig. 6. Diagram showing a state chart of toaster reactive behaviour (left) with a demanded action sequence overlaid, the demanded action sequence (top right), and a picture of the toaster with relevant interface features labelled (bottom right).

- processing stages e.g. perception, cognition and response;
- perceptual modalities e.g. visual, auditory and spatial;
- visual channels e.g. focal vs. ambient;
- processing codes of combined modalities e.g. manual–spatial, vocal-verbal.

A computational multiple resource model is then used to predict performance of the basis of bottlenecks in processing resulting from the competition between cognitive demand elements sharing resource categories during interaction. In line with theories of cognitive interaction and working memory it is assumed that perceptual and motor processing require some cognitive resources (Baddeley, 2000). The four workload dimensions of multiple resource theory could be used for evaluating the demands on the user made by product interactions. These scales are: visual, auditory, cognitive and psychomotor demands (VACP). Considerable research has resulted in the development of an easily deployable quantitative rating scale for each of the four dimensions to be used in assigning demand for task elements (Keller, 2002). Two of the scales are shown in Fig. 7.

Applying the VACP scales to the task analysis derived from the state chart mental model yields a profile of demand for each task step. If the action sequence shown in Fig. 6 represents first time use and the assumption is made that people commonly leave their toaster on a particular heat setting most of the time with the power on; the user has to think about four steps: (1) put bread in, (2) depress slider, (3) wait for bread to cook and (4) remove toast (Fig. 8).

0.0	No Psychomotor Activity
1.0	Speech
2.2	Discrete Actuation (button, toggle, trigger)
2.6	Continuous Adjustive (flight control, sensor control)
4.6	Manipulative
5.8	Discrete Adjustive (rotary, vertical thumbwheel, lever position)
6.5	Symbolic Production (writing)
7.0	Serial Discrete Manipulation (keyboard entries)

0.0	No Cognitive Activity
1.0	Automatic (simple association)
1.2	Alternative Selection
3.7	Sign/Signal Recognition
4.6	Evaluation/Judgment (consider single aspect)
5.3	Encoding/Decoding, Recall
6.8	Evaluation/Judgment (consider several aspects)
7.0	Estimation, Calculation, Conversion

Fig. 7. Quantative rating scales for task elements used by Keller based on the dimensions of Wicken's multiple resource theory.

No	Task Steps	Visual	Auditory	Cognitive	Psychomotor	Total
T1	Place bread in slots	5.4	0.0	1.2	2.6	9.2
T2	Set heat control to desired value	5.4	0.0	1.2	2.6	9.2
T3	Depress slider	3.7	1.0	1	2.2	7.9
T4	Observe toaster and wait for toast	7.0	0.0	4.6	0	11.6
T5	Detect toast done	4.0	4.3	3.7	0	12
T6	Remove bread from toaster	5.4	0.0	0	2.6	8

Fig. 8. The VACP profile for a simple toaster state space analysis with six tasks.

Thus, the overall demand profile suggests that whilst using the toaster involves moderate cognitive demand, this is mainly due to the need to keep track of the state of the bread (T4 and T5). The attentional demands are therefore relatively low although there may be other task distractions in the cooking environment. However, it is worth noting that apart from the toast popping up, the toaster does not signal that the bread has been toasted. The toaster also does not demand a high working memory capacity. 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 they plan through the sequence of actions that will be performed on the toaster. There are also no time demands for task actions that could lead to exceeding the working memory time limit on storage. However, in principle, should the multiple tasking, visual cognitive or other cognitive demands made by a product inhibit performance through overload, this approach could predict the exact interactions that were responsible (Fig. 8). For any given design, the aim would be to reduce the number of actions and decisions that users have to make in order to reach their goals, thus reducing the cognitive demands of planning and working memory.

4.5.1. Representing users mental models

A combined approach could use the VACP scales to calculate the demands predicted during multi-tasking. This would utilise multiple resource theory in conjunction with state space task analysis to accommodate multi-modal interactions and enable the quantification of cognitive demand. However, as has been discussed, mental model research suggests that real mental models of such interactions may be fragmented, partial; shared with the artefacts and could possibly be erroneous (Norman, 1986; O'Malley and Draper, 1992; Payne, 2008). The means that to establish this demand will require performing an analysis of the interaction; not in terms of conventional task analysis, but in terms of elements of the user's actual mental model. Recent work on shared mental models suggest that during cooperative and collaborative work individuals may need to maintain both team knowledge and task knowledge and that performance is related to the overlap of knowledge that is shared. This overlap can be measured using correlational and other metrics of similarity and difference between measured model components (Payne, 2008).

As suggested by Norman (1986), Mieczakowski et al. carried out an analysis examining the requirements for a good representation that was capable of accommodating and combining the system (engineering) model; different users' models and designers' mental models. After eliminating a number of representations that did not meet the criteria, she suggests that the necessary representation may have to make explicit tokens for goals, actions, beliefs and objects (GABO), (Mieczakowski et al., 2010a). The resulting representation and associated computational analysis uses adjacency matrices to calculate the similarities and differences between models. It also utilises human input, through the process of directed local transclusion measuring: isomorphism, commonality, sequence and connectivity (Mieczakowski et al., 2010b). The GABO approach has the potential for development of automated supporting tools that will allow the designer to compare sets of users' mental models with those of the engineering model (Fig. 9) and of the product designer. The engineering or systems model of the toaster used in interaction trials is shown in Fig. 9. This is automatically compared with the user's and designer's mental models, as elicited by standard methods such as protocol analysis, and specified in the same format. Transclusion uses semantic similarity measures to identify common goals, actions, beliefs, and objects across models, enabling further analysis of nodes and connectivity to give a metric of similarity along with specific data regarding model regions of interest.

4.5.2. Cognitive task analysis approaches

A number of other approaches to quantifying cognition in interaction have utilised task analysis in combination with cognitive modelling, in the context of diverse but convergent fields. These include: cognitive systems engineering; the analysis of the development of expertise, and the analysis of the performance of decision makers in naturalistic situations, such as fire-fighting. Many of these approaches have mainstream status in theory and practice (Hoffman and Militello, 2009). For example, cognitive systems engineering (CSE) addresses domains such as human decision making and socio-technical considerations in the design of

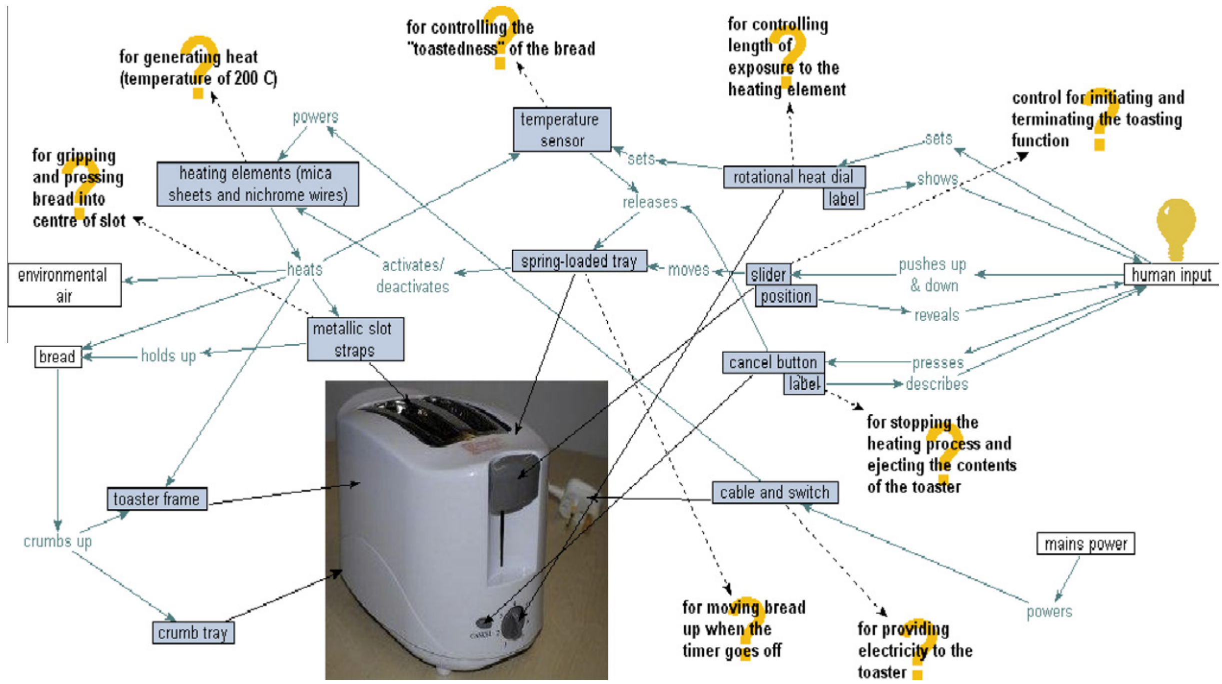

Fig. 9. The GABO model for the engineering structure of the simple toaster.

interfaces for the control of nuclear power stations or air-traffic control. These approaches are frequently concerned with the reduction of errors and failure rates through the analysis of technical work and the working environment. In many cases this involves research into the nature of experts' skills, rules and knowledge and their deployment during work tasks in complex technical systems. For example, designing control interfaces to improve operator's situational awareness during air-traffic control requires the representation of complex safety–critical information in a manner likely to be aligned with human cognitive capabilities. The design and testing of such interfaces requires the performance of complex task analyses, often requiring the accommodation of interactions between different operators and work teams. Success requires achieving a qualitative and quantitative understanding of the relationship between these social–technical features of the controlled systems, the situation, and human cognition (Endsley et al., 2003). These approaches often employ computational modelling of users' cognition using modelling systems such as GOMS and ACT–R, as will be described (Card et al., 1983; Anderson, 2007).

However, there are acknowledged difficulties in current cognitive task analysis (CTA) work, in dealing with the field of design. Design can be seen as the focussing domain of cognitive task analysis in practice (Hoffman and Militello, 2009; Woods, 1998). For example, Hoffman and Militello (2009) identify a number of challenges in applying CTA to design, including the ''leverage problem'' of how cognitive systems analysis can indicate areas of designs where design changes can have a significant impact. This is comparable to the aims of an inclusive design ''exclusion'' audit, where a proposed product design is subjected to capability–demand analysis in order to identify whether it can be used by people with specific quantified impairments. In order to develop system and interface designs that improve general human performance it will also be necessary to go outside of the boundaries of ''normative'' human capability and deal with extra-ordinary human capability in ordinary settings. To date, research underlying CTA approaches has focused on ordinary users in extra-ordinary situations of workload or physical context, such as air-traffic control, piloting modern aircraft, or medical diagnosis. However, as pointed out by Newell (1995), individuals with capability impairments may be faced with the need to interact with designs intended for normal capabilities in everyday situations whilst possessing variable and unusual rages of capability (Newell and Gregor, 1997). Hence, the user of a product may be excluded or disabled from effective use of it by poor design of its interface features, as for example when a keypad is unusable in a low lighting situation.

On this basis, a novel and key requirement criterion of the framework we aim to develop is that it should be capable of accommodating an extra-ordinary range of capabilities, such as those presented by older and disabled people (Newell, 1995, 2006). This means that in any prospective model, representations and processes are required that explicitly address variation across potentially limiting variables such as processing speed, memory and executive function, and perceptual and motor processing. Such approaches and models are rare and have previously been mainly applied to predictions of impaired performance with assistive technology in the context of interfaces designed for disability. For example Horstmann and Levine (1990) applied analytical cognitive modelling techniques based on the GOMS , goals, operators, methods, selection rules approach (Card et al., 1983), to predict the performance of impaired users of augmentative and alternative communication systems (AAC). Such a GOMS approach is based on a hierarchical decomposition of goals and the procedural rules to achieve them, specified as a rule-based knowledge representation. Text entry is an important mode of input for communication purposes for individuals with severe motor impairments resulting in speech and dexterity difficulties. The goal of the modelling was the prediction of text entry speed based on the use of two differing methods computer input technologies. One method used predictive text generation, where the computer offers word completions on the basis of initial character selection using a keyboard layout. This was compared with a switched input selection method, where users activated a switch to indicate their preferred selection while the system cycled sequentially through rows and column of a table display. Predictions based on the GOMS modelling approach were arguably dependant on a number of assumptions. These included assumptions about the strategies employed during user's interactions and the skills necessary to use each technology. There was also a requirement for quantitatively accurate data about important model variables that are necessary in the context of high individual variability that characterises impaired interaction (Newell et al., 1992; Horstmann and Levine, 1992).

5. Criterion 2: identifying predictive variables for modelling

Some data and conclusions are examined here for their implications for measuring the capability–demand relationship. These are taken from research into the relationship between various cognitive measurements or variables and users' performance with products.

5.1. Model variables: candidate empirical measures and model performance

Previous work we have carried out has involved a number of correlation and regression studies of the predictive power of objective and subjective measurements of capability and difficulty. These studies have examined the performance of people over a variety of age groups and capability ranges. In order to validate the effectiveness of this for the framework, an experimental approach was used that compared measures of performance and perceived difficulty with objective measurements of demand for a range of products. This was carried out with a sample of users with varying ranges of capability. These functional capability ranges are characterised by thresholds or human performance envelopes for users with variable or reduced functional capacity. The exclusion or difficulty on a given task was determined by comparing the task demands to the demanded capabilities. To do this it was necessary to arrive at a set of requirements for describing in quantitative terms the interaction between the facets of product demand and the statistics of human capability. This was necessary in order to encompass the scope of inclusive design for the wider capability ranges prevalent in the population, particularly the aging and disabled. Thus, the assessment of compatibility was required to allow for the whole spectrum of capabilities and was based on the distinction between the sensory, motor and cognitive levels of human functioning.

Other researchers have attempted the rigourous modelling of inclusive human performance and tested this by using a detailed quantification of a range of human capabilities (Fleishman and Quaintance, 1984). Perhaps the best example is Kondraske's elemental resources model (ERM) of human performance (Kondraske, 2000). This is based on a calculation of compatibility by utilising resource-demand constructs such as visual acuity, contrast sensitivity, working memory capacity, force generation and movement capabilities, for a set of basic functions. Over a number of studies, Kondraske explored the relationships between high-level measurements of performance and low-level performance resources such as specific perceptual, cognitive and motor abilities. In conclusion, warning that correlations and regressions cannot capture the complexity of such relationships, he proposed a resource economic system performance model based on the likelihood of limitations in specific resource utilisations and the common non-linear thresholds that resulted from combining the measured variables (Kondraske, 2000).

The aim of such approaches has been to find the variables that have the highest correlations to higher level constructs. This is to say, which variables are mostly responsible for performance. This is a conventional approach in human factors and ergonomics for creating predictive models. However, the approach requires large enough numbers of cases to achieve statistical significance for multiple regressions. Persad et al. (2007) carried out an exploratory study looking for evidence of specific qualitative or quantitative relationship between sets of performance and capability measures. They used a number of tasks and products that had the potential to give rise to difficulties in interacting with product features. They also examined how well a number of plausible models of combined performance limitation were able to account for the ability to perform a task.

Four consumer products used in activities of daily living: a clock–radio, a mobile phone, a food blender and a vacuum cleaner, were presented to older and impaired participants whose sensory, cognitive and motor capabilities were evaluated using objective capability tests. These were chosen to represent a range of interface elements, including both informational, technological and computing interfaces (ICT) and control features. Nineteen participants in total participated in the study, ranging from 44 years to 82 years. Of the 19 participants that participated in the study, 12 were female and 7 were male. Participants were selected using a screened, stratified sampling approach in order to represent sufficient qualitative and quantitative variation in capabilities.

First, the users performed a number of randomly assigned standard tasks with each of the products while being videotaped (Fig. 10). The results were analysed to determine how well the capability measures collected prior to task performance predicted the difficulty encountered in actual use, using the capability–demand model. Secondly, a battery of capability tests was administered. These tests included sensory items, such as visual acuity, contrast sensitivity, and hearing level; cognitive tests, such as that of working memory, speed of processing, and long-term memory; and motor tests such as force exertion in various positions, balance, and range of motion. 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.

In general, the mean self-report ratings of difficulty appeared to reflect the relative difficulty of task actions with the products as would be expected. For example, for the mobile phone the cognitively related actions were rated more difficult than perceptual or motion actions in a profile with some face validity (Fig. 11).

This and other results (Persad et al., 2007, 2010) suggested that the approach taken could identify limiting variables that were significant for specific individual's interactions. However, it also showed that specific low-level functional measurement variables alone were poorly correlated with overall performance. The relationship between a number of variables for ICT (mobile phone) and instrumental activities of daily living (IADL) products (blender, vacuum cleaner) are shown as a typical example of the analysis (Figs. 12 and 13).

Persad et al. (2010) also tested several cognitive capability models relating four measured variables: (1) short term working memory; (2) visual–spatial working memory; (3) speed of processing, and (4) long-term memory (Table 1).

In general, over perceptual and motion variables, there was a marked contrast between low-level and combined capability measures. Some low-level capabilities such as log contrast sensitivity for a specific text size range for screen text and button labels were apparently related to specific measures of rated difficulty for specific visual tasks (Fig. 12). Contrast sensitivity tests are measures of the threshold at which elements of contrasting grey can be detected for different sizes of image elements (such as text). Each graph plots the contrast capabilities of participants at a particular element size measured in log units (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.

Fig. 10. Users performing everyday tasks with the range of products trialed.

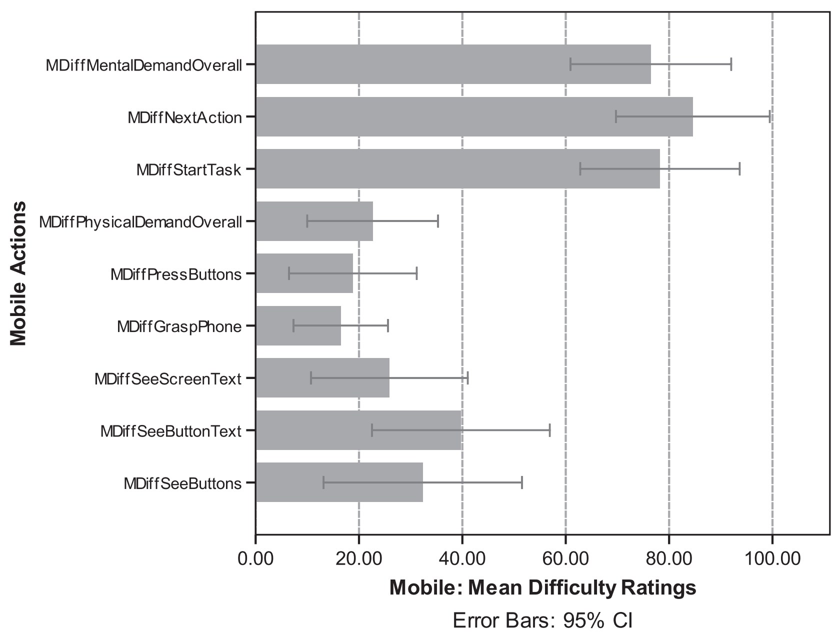

Fig. 11. Mean difficulty ratings across tasks for the mobile phone.

A linear model accounted for a significant amount of the variance for actions involving reading textual features on the products. Fig. 12 shows fairly strong negative relationships for seeing the button labels ($r = -0.766$, $p < 0.01$) of the clock radio. Strong relationships were also observed for seeing the screen text on the mobile $r = -0.688$, $p < 0.01$). However, this was not the case for seeing most product parts with the ICT technology and in general this was not the case for numerous individual low-level measures (Fig. 13, top) that were poorly correlated with perceived task demand.

Interestingly, a number of high level and compound measurements combining several multiple low-level measures proved to correlate well with general performance measures. For example, frustration and overall mental and physical demand measures were clearly related. A linear model accounted for significant amounts of the variance of frustration with physical demand ($r = 0.707$, $F = 19.7$, $df = 18$, $p < 0.001$ and mental demand: $r = 0.547$, $F = 20.5$, $df = 18$, $p < 0.001$). Errors were also well predicted by the Euclidian cognitive capability model (CogCapModel-Euclid) for two different products (Fig. 13, bottom). This could be partially due to the influence of the visual–spatial working memory measure, that showed high correlations with blender errors ($r = -0.711$, $p < 0.01$ and vacuum cleaner errors $r = -0.804$, $p < 0.01$).

This suggests that the more general performance measure of errors during the task, were partially accounted for by a cognitive model based on the cognitive capability–demand relationship. This may be because overall human ability is the result of a combined performance such that the contribution of the interacting

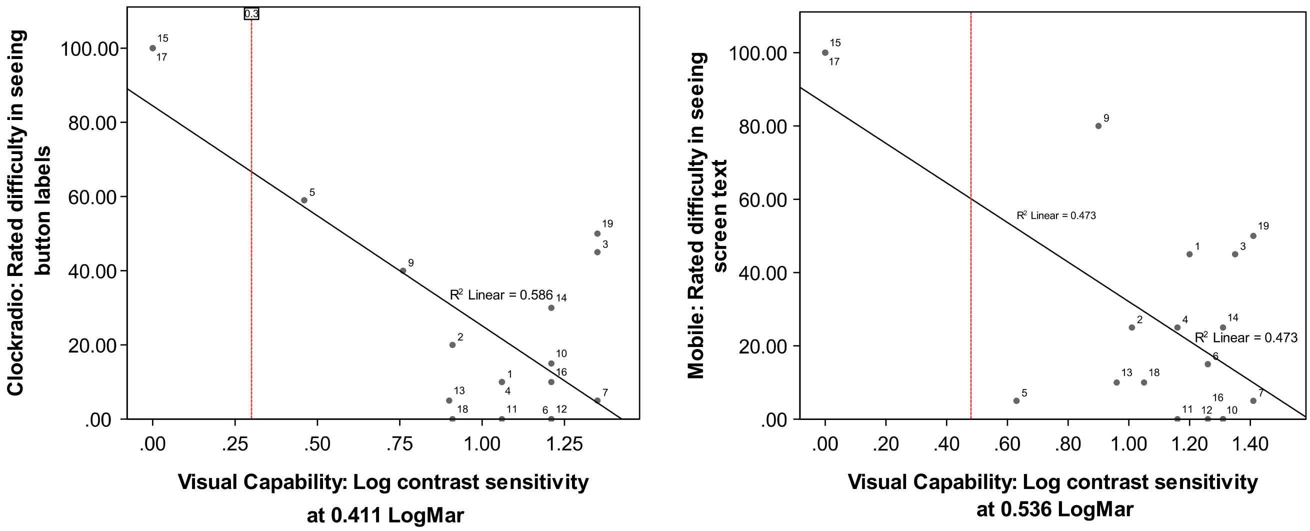

Fig. 12. Log contrast sensitivity for a specific text size range for screen text and button labels and rated difficulty (vertical line indicates required capability).

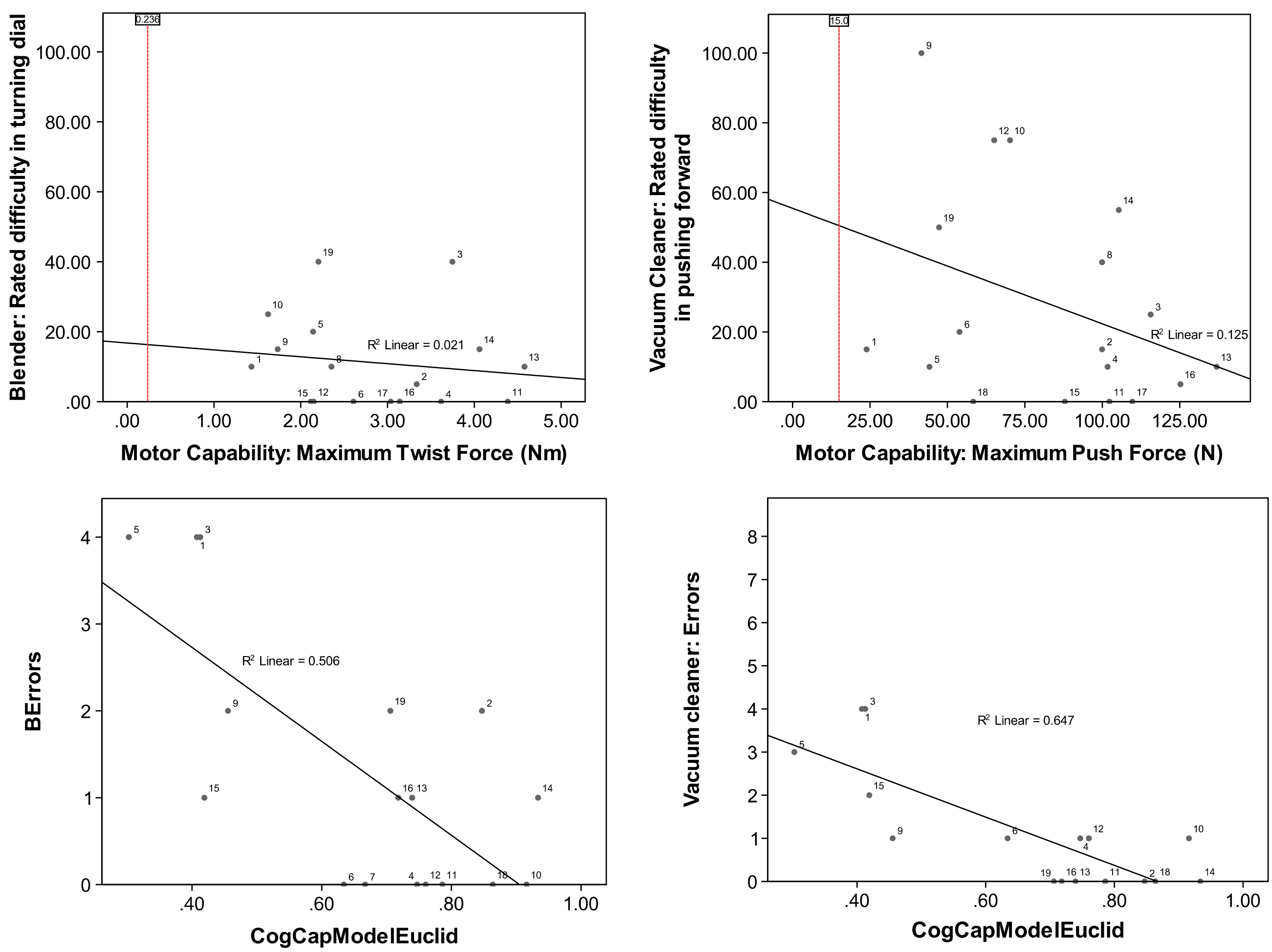

Fig. 13. Low level measures such as twist and push forces were not well related to rated difficulty (TL,TR) whereas more general cognitive measures, such as number of errors were well accounted for by a Euclidian combined cognitive model for the blender (L) and vacuum cleaner (BR).

underlying capabilities are varied according to the situation as demanded by context, task and specific product features. A predictive model might therefore take the form of a combined capability score based on how the user can compensate for some specific capability weakness by employing compensatory strategies that utilise other strengths. Alternatively, the individual may be able to use capabilities in combination in an optimal holistic way that maximises success for specific tasks. In this case it might be that a person whose individual physical measurement suggested that they would be unable to perform a specific action, could in fact

Table 1
Alternative cognitive capability models relating four measured variables.

Model	Description	Equation
Max	The maximum of the four cognitive variables	Max (a–d)
Min	The minimum of the four cognitive variables	Max (a–d)
Euclidean	The euclidean metric is the square root of the sum of the squares of the four cognitive variables	$(a^2 + b^2 + c^2 + d^2)^{1/2}$

modify their physical motion and actions in order to utilise other muscle groups in concert to develop the force required to achieve the same goal. In the case of perceptual capability, a visual weakness could be compensated for by use of multimodal channels, such as hearing, touch and body-sense). Cognitive impairment could be supported by learnt meta-cognitive strategies about the individual's own cognitive capabilities. In this way the user may utilise knowledge of their own cognitive limitations to optimise performance (Anderson, 2007). Alternatively, the user could place greater reliance on well-established crystallised memories or general reasoning using declarative and general procedural knowledge stored in mental models.

However, these results have only indicative status because of the small sample size and restricted selection of measurements. Continuing work is currently examining 100 individuals aged between 50 and 65, on 170 objective and subjective, physical, social and psychological measures. The aim of this will be to derive candidate predictive factors and variables using multiple regression and principle component analysis. These will then be trialed in analytical inclusive design evaluation and the resulting findings can also be used in testing candidate cognitive user models for their ability to simulate the predictions and to accommodate learning, workload and prior experience.

6. Criterion 3: prior experience and learning in the context of ageing

The fields of interaction design and usability have focussed on instantaneous interaction, but the effects of prior experience are evidently important. Extant theories debate the nature of mental models or knowledge structures and their content but less emphasis has been given to the effects of the various contributors to "unconscious" prior experience and their role as part of the ageing individual's capability during real-time interaction with a product. A number of experiments have examined the role of prior experience, age and cognitive capability in individuals' performance with daily living products (Langdon et al., 2007, 2010; Blackler et al., 2003; Blackler, 2006). For example, two microwave ovens were tested that had the same underlying functionality, but with the interface variations of dial or button control (Lewis et al., 2008). The differences in performance were such that dials were found to be easier to use for both younger users and also those with higher cognitive ability. This was not related to prior experience as measured in a product knowledge questionnaire.

However, it was possible that users possessed some degree of prior experience with specific interface elements and their use. Hence, in a second training transfer experiment with digital analogue broadcasting (DAB) radios (Fig. 14), participants were trained to criterion with a common base product to investigate the performance impact that resulted from switching to one of two different interfaces that varied in known interface properties (Langdon et al., 2010). Manipulating the differences between the features of the different interfaces led to varying difficulties in learning and use. In particular, times and errors performance measures were compared for the use of a set of two interfaces that differed in known feature dimensions. However, the extent to which prior experience was a result of generational or specific product feature familiarity remained unclear from the results, as factors such as age and cognitive capability may have interacted to give the performance obtained.

The performance of the users during the trials in terms of time to complete standard tasks and number of cognitive errors made was expected to be dependent on at least four potential components:

- prior experience;
- age-related capability decline;
- generational effects;
- knowledge-based strategies.

In addition, it was expected that performance would be affected by individual task and product difficulty, as well as by the use of speed–accuracy trade-off strategies and the differential cognitive capabilities of the users. Forty-two participants were divided between six conditions in a three-factor, $2 \times 3 \times 3$ mixed-measures design using the factors: Age (40–60; 60–80), cognitive capability (<95; 95–105; >105), and training trials block. The products were chosen to be similar enough that they varied only in terms of visual features and sequence of activity, rather than functionality. Cognitive capability was measured using a combined scales capability score for age 20 (CCS20). Hence these scores were uncorrected for age and made up from individual sub-scales including: verbal ability, mathematical ability, spatial, logic, pattern recognition, general knowledge, STM, and visualisation and classification (Gardner and Moran, 2006).

Fig. 14. DAB radios: Roberts 1 (L), Roberts 2 (M) and Revo-Uno (R).

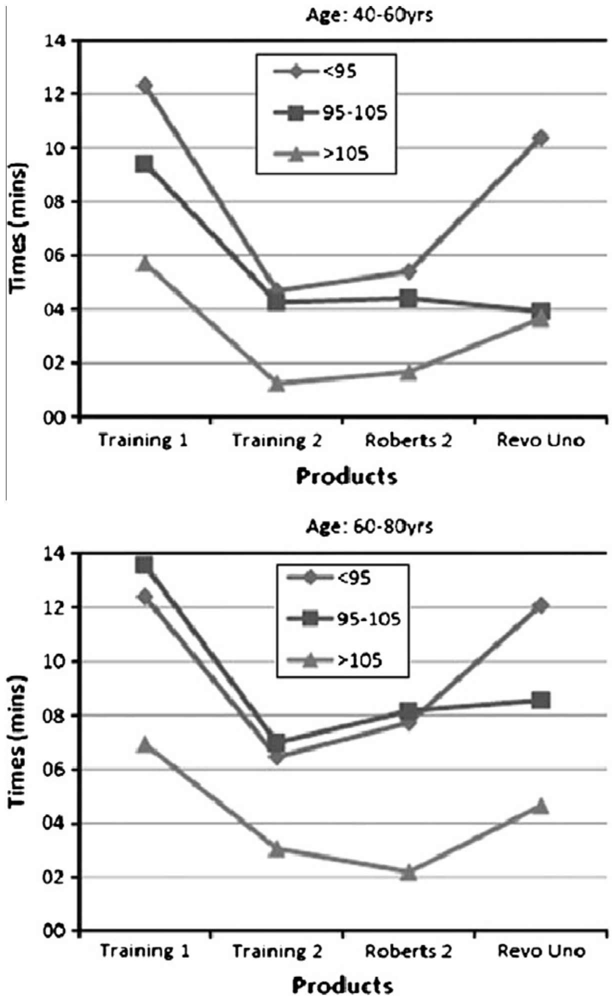

Fig. 15. ANOVA interaction diagrams for task times, three cognitive levels, two age groups, and four trial blocks.

It was found that learning during the training phase occurred quickly over only a few trials (Fig. 15). There were greater variability in response times on the initial untrained trials, and this was presumed to reflect errors resulting from either trial and error approaches or an inappropriate application of prior experience. Training rapidly improved performance with a reduction of inappropriate control actions and the acquisition of sequences that achieved the task goals.

There was a significant effect of age, such that the 40–60 year age group performed better over all cognitive levels and products than the 60–80 year old group in their task completion times ($F(1, 23) = 13.87$, $p < 0.001$). There was also a strong effect of cognitive capability, such that there were significant differences between low, medium and high capabilities across all ages and training trials ($F(2, 23) = 30.32$. $p < 0.001$). As can be seen clearly in the interaction diagrams (Fig. 15), the higher cognitive capability group times were faster than the other groups for nearly all combinations of product age and training. There was a strong effect of training trial block on times to complete tasks overall ($F(3, 23) = 20.024$, $p < 0.001$), such that the first "Training 1" trial was significantly different from all other trials (Bonferoni $p < 0.01$). The "Revo-Uno" times were significantly different from all other trials at $p < 0.05$, and the "Roberts 2" trial was significantly different from the "Revo-Uno" at $p < 0.05$. Therefore, transfer to a second product (Roberts 2) with virtually identical underlying functionality and interface sequences but different layout and appearance of controls, gave rise to a significant increase in time to complete set tasks, showing a strong learning effect of generic functionality consistent with both rule-based and knowledge-based learning. However, error analysis also suggested that small variations of interface function and appearance led to a time consuming misperception and re-exploring of interface functionality. The third product (Revo-Uno) had virtually identical underlying functionality but controls that were very different in usage, layout and appearance. This affected performance by doubling the times and errors. Error analysis suggested that this was due to the misapplication of sequences learnt with the training radio and the resulting performance of actions necessary to explore differences in function through observation of action outcomes. However, the fact that the training trials evidently reduced some of the impact of the different interface of the Revo on performance, suggests that some limited degree of learning of product functionality other than of purely procedural use of the interface took place. This would correspond to the development of a knowledge-based, declarative mental model.

6.1. Prior experience: implications for measuring cognitive demand

These training transfer results are consistent with a general reduction of performance with different interfaces due to age and cognitive capability and also suggest that users were able to learn the new interface sequences more effectively when they possessed higher capability. The learning of new interface sequences was very rapid and correlational analysis suggested that it was dependent on working memory and instantaneous interaction capability. It may also have been the case that users were learning procedural sequences of actions rapidly and effectively, but were also forming declarative models of the interface. However, the evidence further suggested that this was primarily the result of a random selection or means-end analysis as described by Anderson (2007). This interpretation would be consistent with literature on the learning of skilled performance in human factors tasks, and the analysis of human error proposed by Reason (1990). It is also consistent with Freudenthal's (1998) model of interaction learning for older age groups that proposes an additional level of "condition–action" interaction in situations where knowledge-base generic schemas are lacking or inadequate. Blackler's (2006) interpretation of her experimental findings stresses the importance of unconscious experience in "intuitive" interaction with products and deprecates the role of knowledge-based interaction. Observations from these and other studies of the effects of prior experience do not readily support an interpretation in which the majority of users formed coherent, specific mental models of the underlying function of the radios or microwave ovens, from which they were able to reason or hypothesise about functions. Instead, they are more consistent with the notion of fragmented and partial mental model representation that is incomplete without the part played by the product and its behaviour. In this respect they correspond to what is required in order to interact in a goal-directed way with an artefact where interaction is classed as a distributed dialogue between the user and the product (O'Malley and Draper, 1992; Payne, 1992, 2008).

Quantifying the declarative and semantic model of the knowledge in this representation and the way in which procedural rule-based knowledge is acquired and modified during learning that takes place with product use, would require a powerful cognitive user modelling architecture. This model should be capable of predicting performance timings and errors for a complete artefact, taking into account perceptual, cognitive and motor systems. A good candidate for such an approach is Anderson's ACT–R (Anderson, 2007).

7. Cognitive user modelling of the demand–capability relationship

Detailed interaction of demand with variable levels of reduced cognitive capability in extra-ordinary interactions (Newell, 1995) may be quantifiable using simplified approaches tested using user modelling techniques such as GOMS and ACT/R (Anderson, 2007). This could give rise to a better understanding of the design principles applicable to designing interfaces for variable cognitive capability. The Adapted Character of Thought–Rational theory (ACT–R) characterises complex cognition in thought as an interaction between procedural knowledge, regarding how to do something, with declarative knowledge of what things are. Hence, objects in the environment are encoded as representational chunks and operations that can be carried out on them are encoded as transformation or production rules. These response-tendency rules are activated on the basis of their preconditions resulting in changes to the input to working memory in a manner that simulates goal-directed, means-end planning. These operations of working memory in interaction are modelled as the interaction between sensory, visual or auditory processing modules. These interact with their associated storage buffers and the procedural selection, processing, and transfer of information to the response systems for vocal and manual output modules using a limited capacity implementation. Importantly, the procedural rule system selects appropriate declarative, chunked data on the basis of statistical similarity properties of the situation that have been encoded during prior experience of the context. The ACT–R modelling approach has been applied successfully to the prediction performance in many domains (Fleetwood and Byrne, 2002; Byrne, 2001). These include: mathematical reasoning; the design of visual and sequential properties of software interfaces; the role of attentional factors in task execution, and the timing and error performance over a wide range of human cognitive capabilities. Significantly, for the purpose of modelling inclusive cognitive interaction, it has been shown to be capable of accurately modelling erroneous thinking and the characterisation of performance parameters in the transfer from novice to expert, and from exploratory to automatic skilled performance (Anderson, 1982, 2007; Anderson et al., 1992, 2004). Furthermore, the ACT–R architecture's cognitive parameters can be systematically varied to assume various cognitive capability levels in the simulated user. These can be extended to levels found in the wider population that are covered by inclusive design and matched to data from real individuals or values from experimental literature.

7.1. Role of user models

Models such as ACT–R will be essential to address the effectiveness of the proposed inclusive models for quantification as predictive tools for analytical exclusion auditing of alternative designs for inclusive populations. It is acknowledged that such an approach may not capture unique or idiosyncratic variation in capability due to other aspects of ageing or impairment. For example, as has been already mentioned, Newell et al. (1992) argue that it may not be possible using modelling to distinguishing the relative effectiveness of two AAC input interfaces. It may not be possible to operationalise all the user's responses during an interaction although modelling can be used to instantiate various alternatives. The value of the quantitative accuracy in cognitive modelling may lie in its generality. This is not conducive to modelling the high variability found in impairment in the wider population but models could be varied to characterise different types of user. Specific or extreme users may require specially structured models and these can be accommodated by the modelling approach. Complex architectures such as ACT–R are capable of modelling errors, parallel processing and ranges and clusters of performance capability (Horstmann and Levine, 1992; Anderson, 2007). On the other hand, insights from the validation of the modelling process may yield better models of impaired interaction, particularly if cognitive impairment is considered explicitly. It is unclear whether alternative approaches to theorising about individual capability and performance would be more accurate or successful. The unknown or additional unconsidered factors such as specific cognitive impairments, social and emotional factors or communication disorders may affect the ways in which demand from design features can exclude users. These are not predicted by conventional user models (Newell et al., 1992). However, in principle, these factors would be made more salient by the qualitative and quantitative mismatch between the model set-up, the specific predictions made and the observed user behaviour during experimental product trials with ageing and impaired users. Insights gained from these mismatches would then be used to develop better models.

8. Summary, discussion and conclusions

We have aimed to establish the basic elements that are necessary for an empirical and theoretical groundwork for a framework for quantifying inclusive cognitive interaction. These basic elements were required to populate the framework intended for developing research into predicting cognitive capability–demand relationships for analytical inclusive design. The goal has been to make explicit the framework; its antecedents in research and literature from an eclectic, multi-disciplinary background, and to document this research strategy and some exemplars of empirical work derived from it. In the following two sections, we first summarise and discuss the effectiveness with which the framework developed so far has addressed the key issues and problems that were identified in cognitive analytical inclusive design. Secondly, we revisit the framework and assess the extent to which the elements that have been developed have satisfied the original requirements in the light of the derived experiments and research.

8.1. Applicability to key issues in analytical inclusive design

Firstly, we addressed the main theoretical Issues, problems and gaps in knowledge that arise in assessing the cognitive demand of products (Fig. 16). There was a requirement for an adequate underlying model of cognition and interaction to accommodate the breadth of cognition involved in product interaction for inclusive design. In order to populate a predictive model of inclusive cognitive interaction, a set of variables and measurements were required that could be identified as having the power to predict the relationship between changes in demand made by specific product features, such as text size, handle grip dimensions and menu depth; and the resulting exclusion or difficulty arising from capability, such as visual acuity, contrast sensitivity, range of motion and grip strength. The prediction of performance was thought likely to be ineffective without reference to the way in which the prior experience and learning that had taken place with similar interfaces could affect users of varying cognitive capabilities. It was clear that a comprehensive and quantitative user modelling system would be required due to the complexity of interaction between learning, memory and executive functional capacity, and perception and movement. The initial gap analysis generated a list of component elements that were necessary for a research framework and these were then restated as requirements criteria for measuring cognitive demand for inclusive design. To be able to account for the demand arising from moment-by-moment interaction with product features a candidate user model was required that could, in principle, account for performance involving

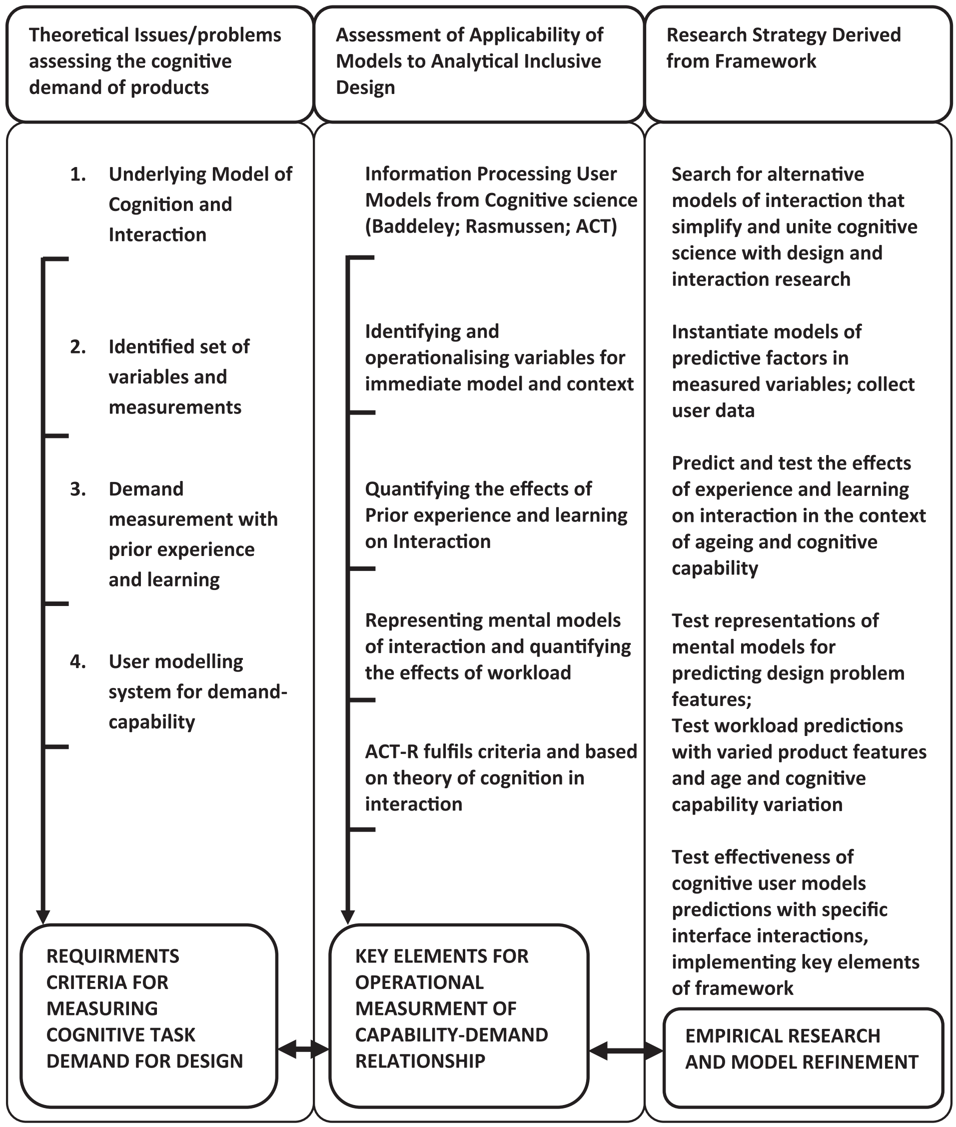

Fig. 16. Table showing the relationship between requirements, key issues and the derived research strategy in the proposed framework.

perceptual, cognitive and movement components of interaction. Such a model would need to account for perceptual phenomena, such as attentional focus and visual search, as well as capacity limitations in the timing and amount of information that could be processed in short term working memory.

Because of the requirement to accommodate prior experience and the similarity of cues in an interaction to previously learnt long-term memories, the framework also needed to distinguish episodic, procedural and declarative components of memory. It was assumed that such memories and their organisation in LTM give rise to mental models in WM that act as internal representations of the task and its context and incorporate the skill-based, rule-based and knowledge-based unconsciously stored information that is necessary for interaction. The framework was required to be capable of quantifying the acquisition of such mental information during interaction as a process of associative learning. A further requirement was that the derived cognitive model should be able to account for both the serial processing of information in different modalities (e.g. sound and visual cues) and also account for processing and workload during parallel sub-tasks, for example, when pressing a key and simultaneously selecting a menu by pointing.

A well-accepted model of cognitive information processing; that of Baddeley (2000), met the criterion for a suitable underlying theoretical architecture for working memory and memory function during product interaction. This model was also consistent with other theories accounting for cognitive workload during interaction that have been developed during human factors modelling of high-performance applications such as piloting aircraft and air-traffic control (Hoffman and Militello, 2009). A combined approach using VACP scales for multi-tasking, and multiple resource theory to accommodate multi-modal interactions was found to have the potential for enabling the quantification of cognitive demand (Keller, 2002; Wickens, 2002). The means to establish this demand would require performing an analysis of the interaction; not using conventional task analysis, but in terms of elements of the user's mental model during interaction. We presented some research on possible representations for combining the system; users' and designer's mental models that suggests that the necessary representation may have to make explicit goals, actions, beliefs and objects and their associated interactions (Mieczakowski et al., 2010b). This approach contrasts, for example, with that of the application of cognitive task analysis to the study of reasoning in expert performance (Hoffman and Militello, 2009).

The necessity for modelling the acquisition of new knowledge regarding a product and the process of integrating that knowledge with other information acquired during prior experience to form skilled, expert behaviour, is within the domain of user models that accommodate learning such as ACT/R. However, the nature of the knowledge acquired and how it can be experimentally distinguished from skill level or rule-based procedural performance during novel product learning could also be investigated for differing age groups and product demands. The results of doing this suggest that generation-specific differences in the performance and learning of novel product interfaces can be demonstrated and that these are related to age and cognitive capability. However, it is also possible to show that, at the same time, infelicities in product design can produce features that are detrimental to all age and capability groups (Wilkinson et al., 2009, 2010).

The effect of specific prior experience of products and product features for different generations and how this affects performance on tasks for varying ages and varying levels of cognitive capability has been addressed by a number of researchers (Langdon et al., 2007, 2010; Blackler et al., 2003; Blackler, 2006; Hurtienne et al., 2008). Hurtienne, et al. established that prior experience has at least two components. Firstly, that of amount of exposure to an interface type and secondly the degree of competence, in terms of skill levels, that the user had acquired. The age-related findings of Blackler (2006) and Langdon et al., 2007, 2010) suggest that performance is affected by general decline in capability due to ageing but is also subject to generational effects and may additionally be affected by learning capability and general cognitive ability. These latter studies, in particular, present considerable evidence that ageing users engage in trial and error or means-end analysis when encountering specific tasks and product features outside of their prior experience. As a principal finding of these applied design studies, this suggests that in the case of concrete rather than abstract tasks users respond using procedural knowledge direct from experience. However, a lack of cues in the situation and environment apparently prompts a knowledge-based exploration of the problem that utilises a declarative and semantic mental model (Anderson et al., 1992). Anderson et al. (1992) tested participants' use of naive physics in anticipating the movements of a falling ball, contrasting predictions of a complete semantic model approach against a "Knowledge in Pieces" account of knowledge primitives extracted directly from experience (diSessa, 1992). The results suggested that novices were unable to utilise predictively a full semantic mental model of the situation of the sort that experts were able to reason with. Instead, they appeared to rely on partial fragments of knowledge that operated explanatorily at the level of salient observable features. Mental models therefore appear to be fragmentary and conflicted when considered as knowledge representations revealed by experiment but they show coherence and utility when seen in the context of a specific product interaction (O'Malley and Draper, 1992). This interpretation of the psychological findings regarding mental models appears to be consistent with other approaches to novice and expert performance discussed above (Rasmussen, 1986; Arnold and Mettau, 2006). It is also taken into account by cognitive modelling approaches such as ACT–R that propose both procedural and declarative memory structures that are coordinated by working memory during the course of cognition (Newell, 1994; Anderson et al., 2004). On this basis, the use of this form of cognitive user model is justified as a response to the framework criteria.

8.2. Conclusions: the framework derived research strategy

We revisit the proposed framework and assess the extent to which the elements that have been developed have satisfied the original requirements in the light of the derived experiments and research. Fig. 16 shows a table based on the strategy adopted for development of the framework. The purpose of the framework was that of directing a research agenda to address the quantification of cognitive demand and capability measurement for the purpose of cognitive analytical inclusive design. The figure shows how the original requirements analysis led directly to: (1) the identification and selection of suitable models of cognition in interaction; (2) research on the effects of prior experience, and (3) variables and user modelling techniques for testing the measurement of capability and its interaction with the demand made by product features. The third column lists the empirical research derived from these key elements and therefore enables an assessment of progress to date (Fig. 16). We have described how the analysis of key elements has led to the identification of suitable research areas. In particular, it has enabled specific studies into:

- the identification of predictive variables for the collection of new data sets;
- the measurement of the effect of prior experience on the learning and use of specific product designs for different age groups and cognitive capability levels;
- the measurement of the relationship between collective users mental models and those of designers;
- the relationship between cognitive workload in the interaction with product features and cognitive capabilities;
- the user modelling of cognitive capabilities and their interactions with specific product features, such as menu structures and the perceptual and physical positioning of controls.

The outcomes of these areas of research, to date, lie principally within the locus of improved understanding of the research questions to be addressed. Areas where the framework has proved to be unable to contextualise research have yet to be identified. It remains to be seen how the insights gained will be ultimately transferred to product design in order to benefit the end-users. However, it is already clear that the development of tools and techniques based on the model will be capable of generating feedback concerning the effectiveness of prediction and therefore will be essential when employed in modifying the elements of the framework. In turn, this will lead to improvement of the requirements criteria in order to increase our understanding of cognitive interaction for inclusive design.

8.3. Some special considerations of impact on the end-user

The principal difference between conventional user modelling and modelling for inclusive demand is the necessity to deal with extra-ordinary ranges of perceptual, cognitive and movement capability in the prediction of performance with product interfaces. It is further necessary to use the theoretical framework in order to develop practical tools and techniques for performing the analytical analysis during the design process. We have proposed a systematic set of criteria for what would constitute an adequate framework for modelling inclusive cognitive demand. Taken together, the considerations of model, workload, skilled performance, good representation of knowledge, and empirical studies of interaction by participants in differing age and capability groups, has generated profitable research outcomes. This combined approach taken to cognitive demand has shown itself to be capable of satisfying the proposed criteria and furthermore has demonstrated that it has the potential to be utilised in practice to identify important predictive variables and relate these to an underlying model of interaction.

An important consequence of successful development of the framework and development of design tools based on the sort of cognitive model intended, will be that designers need not have a detailed understanding of cognitive science in order to find out where aspects of their designs may exclude users with a certain level of cognitive capability. Nor, it is intended, will it be necessary for them to carry out expensive and time-consuming user trials with numerous participants. However, the importance of relating the model to the knowledge and information that comes from psychology, the clinical and therapeutic disciplines, and design practice, cannot be underestimated. These will need to be incorporated into any tools developed for analytical inclusive design. Hence, user information and indexed knowledge, in the form of personas, films and theatrical presentation of key design and user capability issues, will be required to contextualise the design analysis. In particular, such analytical design tools should carry information regarding the importance of variability in capability, within individuals and within the population, in order to capture the diversity of older and impaired populations. Information about the social and economic context of interaction should also be available and indexed to the quantitative analysis. Tools should provide clear information regarding critical issues such as the relationship between impairment and disability in given environments; ethical issues in the collection and recruitment of users in design trials; the importance of informed consent, and the need for provision of adaptation into designs (Newell and Gregor, 1997; Newell et al., 2008).

A number of tools, techniques and methods have been generated by related work on assessing the perceptual and physical demand of products. For example, these have developed the use of psychological judgements for rating the demand made by particular product features on scales such as visual capability that correspond to known levels of population capability in survey data (Waller et al., 2010b). Currently, the present framework is also contributing to the collection of new data sets. These will be representative of the UK population and relate to specific items derived from the framework, such as contrast sensitivity, memory, executive function, and speed of processing. In conjunction with this, multiple regressions of over 100 perceptual, cognitive and physical variables and performance with a number of everyday products is expected to suggest predictive factors and their combination in interaction that will be candidates for instantiation in future user modelling.

Research based on the developing framework has already been used with some success in practice. For example, the GABO representation for mental models of daily living products is being tested in an industrial context as a tool for identifying areas of disagreement between user's models of a design and the designer's intent (Mieczakowski et al., 2010a). This application has already identified critical further areas for research, while feedback from the studies indicates that there may be value in application to other areas, such as engineering design. Such research tests the predictive power of theories, and where this is low, the resulting analysis of areas where model components are not predictive over the required parameter ranges will lead directly to stronger models or improved model requirements. Ultimately, this should lead to a better conceptual understanding of the capability of the human system to accommodate multiple interacting impairments. A novel question already illuminated by the developing framework, will be how it is that an individual's cognitive capabilities can be used to compensate for perceptual and movement impairments and whether a general underlying mechanism is involved in this compensatory response that may be operationalised, modelled and investigated.

Ultimately, the utility of the framework will be judged by experimental tests of the accuracy of prediction of exclusion from demand and capability. Further research is underway to address this strategy in practice and to test whether designers can create inclusive designs for ICT products using methods and tools based on considerations derived from the combined elements of the approach. In the real world, metrics of validity must depend on whether the resulting design changes lead to better designs of more inclusive products that can be used without special effort by older and impaired users.

Acknowledgements

This work has been funded by the UK Engineering and Physical Sciences Research Council as part of the Inclusive Design 3 Project. We also acknowledge the contribution of a number of researchers and research students including Dave Brown of Worcester Polytechnic Institute, Massachusetts, US; and Tim Lewis, Chris Wilkinson and Anna Mieczakowski.

References

Anderson, J.R., 1982. Acquisition of cognitive skill. Psychological Review 89 (4), 369–406.
Anderson, J.R., 2007. How Can the Human Mind Occur in The Physical Universe? Oxford University Press, New York.
Anderson, T., Tolmie, A., Howe, C., Mayes, T., Mackenzie, M., 1992. Mental models of motion. In: Rogers, Y., Rutherford, A., Bibby, P.A. (Eds.), Models in the Mind – Theory, Perspective and Applications. Academic Press Ltd., London, pp. 56–71.
Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Quin, Y., 2004. An integrated theory of the mind. Psychological Review 111 (4), 1036–1060.
Arnold, A.G., 1999. Mental effort and evaluation of user interfaces: a questionnaire approach. In: Bullinger, H.-J., Ziegler, J. (Eds.), Proceedings of the HCII '99 International Conference on Human–Computer Interaction, vol. 1. Lawrence Erlbaum, Mahwah, NJ, pp. 1003–1007.
Arnold, A.G., Mettau, P., 2006. Action facilitation and desired behaviour. In: Verbeek, P.P., Slob, A. (Eds.), User Behavior and Technology Development: Shaping Sustainable Relations between Consumers and Technologies. Springer, pp. 13–20.
Baddeley, A.D., 2000. The episodic buffer: a new component of working memory? Trends in Cognitive Sciences 4 (11).
Blackler, A., 2006. Intuitive Interaction with Complex Artefacts. PhD Thesis, Queensland University of Technology, Brisbane.
Blackler, A., Popovic, V., Mahar, D., 2003. The nature of intuitive use of products: an experimental approach. Design Studies 24 (6), 491–506.
Byrne, M.D., 2001. ACT–R/PM and menu selection: applying a cognitive architecture to HCI. International Journal of Human–Computer Studies 55, 1–4.
Card, S., Moran, T.P., Newell, A., 1983. The Psychology of Human–Computer Interaction. Erlbaum, Hillsdale, NJ.
Carroll, J.M., Carrithers, C., 1984. Blocking learner error states in a training wheels system. Human Factors 26 (4), 377–389.
Clarkson, P.J., Coleman, R., Keates, S., Lebbon, C. (Eds.), 2003. Inclusive Design: Design for the Whole Population. Springer-Verlag, London.
Coleman, R., 2001. Designing for our future selves. In: Preiser, W.F.E., Ostroff, E. (Eds.), Universal Design Handbook. MacGraw-Hill, New York, pp. 41–425.
DFS-SN4090, 1997. Disability Follow-up to the 1996/97 Family Resources Survey. <http://www.esds.ac.uk/findingData/snDescription.asp?sn=4090>. (accessed 10.06.10).
diSessa, A., 1992. Phenomenology and the evolution of intuition. In: Rogers, Y., Rutherford, A., Bibby, P.A. (Eds.), Models in the Mind – Theory, Perspective and Applications. Academic Press, London, pp. 15–33.
Docampo-Rama, R.M., 2001. Technology Generations Handling Complex User Interfaces. PhD Thesis, TU Eindhoven. ISBN 90-386-0913-2.
Endsley, M.R., Bolte, B., Jones, D.G., 2003. Designing for Situation Awareness: An Approach to User Centred Design. Taylor & Francis.
Fleetwood, M.D., Byrne, M.D., 2002. Modelling icon search in ACT–R/PM. Cognitive Systems Research 3 (1), 25–33.
Fleishman, E.A., Quaintance, M.K., 1984. Taxonomies of Human Performance. AP, Orlando.
Freudenthal, T.D., 1998. Learning to Use Interactive Devices; Age Differences in the Reasoning Process. PhD Thesis, Eindhoven University of Technology.
Freudenthal, A., 1999. The Design of Home Appliances for Young and Old Consumers. Series Ageing and Ergonomics, Part 2. PhD Thesis, Delft University Press, The Netherlands. ISBN 90-407-1755-9.
Gardner, H., Moran, S., 2006. The science of multiple intelligences theory: a response to Lynn Waterhouse. Educational Psychologist 41 (4), 227–232 (Fall 2006).
Gentner, D., Stevens, A.L. (Eds.), 1983. Mental Models. LEA.
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.
Hart, S.G., Staveland, L.E., 1988. Development of a multi-dimensional workload rating scale: results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. Elsevier, Amsterdam, The Netherlands, pp. 139–183.

Hoffman, R.R., Militello, L.G., 2009. Perspectives on Cognitive Task Analysis: Historical Origins and Modern Communities of Practice. Psychology Press, Taylor & Francis, NY.

Horstmann, H.M., Levine, S.P., 1990. Modelling of user performance with computer access and augmentative communication systems for handicapped people. Augmentative and Alternative Communication 6, 231–241.

Horstmann, H.M., Levine, S.P., 1992. Modelling AAC User Performance. Response to Newell, Arnott, and Waller. Augmentative and Alternative Communication 8 (2), 92–97.

Hurtienne, J., Weber, K., Blessing, L., 2008. Prior experience and intuitive use: image schemas in user centered design. In: Langdon, P., Clarkson, J., Robinson, P. (Eds.), Designing Inclusive Futures. Springer, London, pp. 107–116.

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 (UAIS) 1, 56–76.

Johnson-Laird, P.N., 1980. Mental models in cognitive science. Cognitive Science 4, 71–115.

Keates, S., Clarkson, P.J., 2004. Countering Design Exclusion – An Introduction to Inclusive Design. Springer-Verlag, London, UK.

Keller, J., 2002. Human performance modelling for discrete-event simulation: workload. In: Yücesan, E., Chen, C.-H., Snowdon, J.L., Charnes, J.M. (Eds.), Proceedings of the 34th IEEE 2002 Winter Simulation Conference, San Diego.

Kieras, D.E., Bovair, S., 1984. The role of a mental model in learning to operate a device. Cognitive Science 8, 255–273.

Kondraske, G.V., 2000. A working model for human system-task interfaces. In: Bronzino, J.D. (Ed.), 2nd ed., The Biomedical Engineering Handbook, vol. 2 CRC Press. pp. 147-141 – 147-118.

Langdon, P.M., 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.

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 (ISSN:1615-5289).

Lewis, T., Clarkson, P.J. 2005. A user study into customising for inclusive design. In: Proceedings of Include 2005, London, UK.

Lewis, T., Langdon, P.M., Clarkson, P.J., 2008. Prior experience of domestic microwave cooker interfaces: a user study. In: Langdon, P.M., Clarkson, P.J., Robinson, P. (Eds.), Designing Inclusive Futures. Springer, London, pp. 95–106.

Lintern, G., Roscoe, S.N., Sivier, J., 1990. Display principles, control dynamics, and environmental factors in pilot performance and transfer of training. Human Factors 32, 299–317.

Marr, D., 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. W.H. Freeman, San Francisco.

Martin, J., Elliot, D., 1992. Creating an overall measure of severity of disability for the office of population census and surveys disability survey. Journal of the Royal Statistical Society Series A 155 (1), 121–140.

Mieczakowski, A., Langdon, P.M., Clarkson, P.J., 2010a. Investigating designers' cognitive representations for inclusive interaction between products and users. In: Langdon, P.M., Clarkson, P.J., Robinson, P. (Eds.), Designing Inclusive Interactions. Springer, London, pp. 133–145. ISBN 9-781849-961653.

Mieczakowski, A., Langdon, P.M., Bracewell, R., Clarkson, P.J., 2010b. Towards a model of product–user interaction: a new data modelling approach for designers. In: International Design Conference – Design 2010, Dubrovnik, Croatia, May 17–20, 2010.

Newell, A., 1994. Unified Theories of Cognition: The William James Lectures. Harvard University Press. ISBN 0-674-92101-1.

Newell, A.F., 1995. Extra-ordinary human computer operation. In: Edwards, A.D.N. (Ed.), Extra-ordinary Human–Computer Interactions. Cambridge University Press.

Newell, A.F., 2006. Older people as a focus for inclusive design. Gerontechnology 4 (4), 190–199.

Newell, A.F., Gregor, P., 1997. Human computer interfaces for people with disabilities. In: Helander, M., Landauer, T.K., Prabhu, P. (Eds.), Handbook of Human Computer Interaction, pp. 813–824. ISBN 0-444-1862-6.

Newell, A.F., Arnott, J.L., Waller, A., 1992. On the validity of user-modelling in AAC: Comments on Horstmann and Levine (1990). Augmentative and Alternative Communication 8, 89–92.

Newell, A.F., Carmichael, A., Gregor, P., Alm, N., Waller, A., 2008. Information technology for cognitive support. In: Sears, A., Jacko, J.A. (Eds.), The Human–Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications Human Factors and Ergonomics, Section 5, Designing for Diversity, 2nd Ed., CDC, Taylor & Francis, pp. 811–882, Chapter 41.

Nichols, T.A., Rogers, W.A., Fisk, A.D., 2006. Design for aging. In: Salvendy, G. (Ed.), 3rd ed., Handbook of Human Factors and Ergonomics Wiley, pp. 1418–1445.

Nicolle, C., Abascal, J.G. (Eds.), 2001. Inclusive Design Guidelines for HCI. Taylor & Francis, London, UK.

Norman, D.A., 1986. Some observations on mental models. In: Gentner, S., Stevens, A.L. (Eds.), 1983. Mental Models. Lawrence Erlbaum Associates, pp. 7–14.

Norman, D.A., 2002. The Design of Everyday Things. Basic Books, Doubleday/ Currency, New York, NY, USA.

O'Malley, C., Draper, S., 1992. Representation and interaction. In: Rogers, Y., Rutherford, A., Bibby, P.A. (Eds.), Models in the Mind – Theory, Perspective and Applications. Academic Press, London, pp. 56–71.

Payne, S.J., 1992. On mental models and cognitive artefacts. In: Rogers, Y., Rutherford, A., Bibby, P. (Eds.), Models in the Mind – Theory, Perspective, and Application. AP, London.

Payne, S.J., 2008. Mental models in human–computer interaction. In: Sears, A., SearsJacko, J.A. (Eds.), 2nd ed., The Human–Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications Human Factors and Ergonomics CDC, Taylor & Francis, pp. 64–76 (Section 3).

Persad, U., Langdon, P., Clarkson, J., 2007. Characterising user capabilities to support inclusive design evaluation. Universal Access in the Information Society 6 (2), 119–135.

Persad, U., Langdon, P., Clarkson, P.J., 2010. Supporting Inclusive Evaluation with User Capability Data. In: Proceedings of the 3rd Applied Human Factors and Ergonomics (AHFE) International Conference, 2010.

Petrie, H., 2001. Accessibility and usability requirements for ICTs for disabled and elderly people: a functional classification approach. In: Nicolle, C., Abascal, J.G. (Eds.), Inclusive Guidelines for Human Computer Interaction. Taylor & Francis, London. 0-748409-48-3.

Porter, J.M., Marshall, R., Freer, M., Case, K., 2004a. SAMMIE: a computer aided ergonomics design tool. In: Delleman, N.J., Haslegrave, C.M., Chaffin, D.B. (Eds.), Working Postures and Movement Tools for Evaluation and Engineering. CRC Press LLC, Boca Raton, pp. 454–462.

Porter, J.M., Case, K., Marshall, R., Gyi, D.E., Sims, R.E., 2004b. 'Beyond Jack and Jill': designing for individuals using HADRIAN. International Journal of Industrial Ergonomics 333 (March), 249–264.

Rabbitt, P.M.A., 1993. Does it all go together when it goes? The Nineteenth Bartlett memorial lecture. The Quarterly Journal of Experimental Psychology 46A (3), 385–434.

Rasmussen, J., 1986. Information Processing and Human–Machine Interaction: An Approach to Cognitive Engineering, North-Holland Series in System Science and Engineering, 12, NY.

Reason, J., 1990. Human Error. Cambridge University Press, UK, NY.

Rogers, Y., Rutherford, A., Bibby, P. (Eds.), 1992. Models in the Mind – Theory, Perspective, and Application. AP, London.

Ryu, H., Monk, A., 2004. Analysing interaction problems with cyclic interaction theory: low-level interaction walkthrough. Psychology Journal 2 (3), 304–330.

Schaie, K.W., 1988. Variability in cognitive functions in the elderly: implications for societal participation. Basic Life Science 43, 191–211.

Story, M.F., Mueller, J.L., Mace, R.L., 1998. The Universal Design File. North Carolina State University, The Centre for Universal Design.

Thimbleby, H., 2007. Press On: Principles of Interaction Programming. MIT Press.

Vanderheiden, G.C., Vanderheiden, K., 1992. Guidelines for the Design of Consumer Products to Increase Their Accessibility to People with Disabilities. Trace R & D Center.

Waller, S.D., Williams, E.Y., Langdon, P.M., Clarkson, P.J., 2010a. Understanding the co-occurrence of ability loss. In: Langdon, P.M., Clarkson, P.J., Robinson, P. (Eds.), Designing Inclusive Interactions. Springer, London. ISBN 9-781849-961653.

Waller, S.D., Williams, E.Y., Langdon, P.M., Clarkson, P.J., 2010b. Quantifying Exclusion for Tasks Related to Product Interaction. In: Langdon, P.M., Clarkson, P.J., Robinson, P. (Eds.), Designing Inclusive Interactions. Springer, London, pp. 57–69. ISBN 9-781849-961653.

Wickens, C.D., 2002. Multiple resources and performance prediction. Theoretical Issues in Ergonomic Science 3 (2), 159–177.

Wickens, C.D., Hollands, J.G., 2000, 3rd ed. Engineering Psychology and Human Performance Prentice-Hall. ISBN 321-04711.

Wilkinson, C., Langdon, P.M., Clarkson, P.J., 2009. Investigating prior experience and product learning through novel interface interaction. In: 5th International Conference on Universal Access in Human–Computer Interaction (Part I), San Diego, California, USA, pp. 654–664.

Wilkinson, C., Langdon, P.M., Clarkson, P.J., 2010. Prior experience and learning: generational effects upon interaction. In: Langdon, P.M., Clarkson, P.J., Robinson, P. (Eds.). Designing Inclusive Interactions. Springer, London, pp. 145–154. ISBN 9-781849-961653.

Woods, D.D., 1998. Designs are hypotheses about how artifacts shape cognition and collaboration. Ergonomics 41, 168–173.