


This item is held in Loughborough University's Institutional Repository (<https://dspace.lboro.ac.uk/>) and was harvested from the British Library's EThOS service (<http://www.ethos.bl.uk/>). It is made available under the following Creative Commons Licence conditions.




C O M M O N S D E E D


**Attribution-NonCommercial-NoDerivs 2.5**

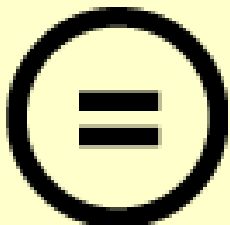
**You are free:**

- to copy, distribute, display, and perform the work

**Under the following conditions:**

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.


 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

**Your fair use and other rights are in no way affected by the above.**

This is a human-readable summary of the [Legal Code \(the full license\)](https://creativecommons.org/licenses/by-nc-nd/2.5/).

[Disclaimer](#) 

For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

## Abstract

---

### **Towards developing and improving effective interaction design tools**

This research began by addressing the question: can effective interface design guidelines be produced for use in the design of future consumer product technologies (CPT)?

A literature review explored published studies evaluating existing Human-Computer Interaction guidelines to establish their effectiveness in relation to CPT. Through this review, effectiveness was found to be limited but potentially could be improved using user-centred design methods. In response, six short studies were undertaken to produce user-centred CPT guidelines and to evaluate them using two sets of effectiveness criteria: specificity and applicability. These studies supported findings from the HCI literature. Despite improving the specificity and applicability of the CPT guidelines, passive, non-bespoke design guidelines have still been shown to have little impact on interaction design activity.

Other links between research and practice needed to be identified. Two further field investigations indicated that, whilst the use of ergonomics methods was limited in commercial design consultancies, certain types of participative methods considering 'situated design in context' might be helpful.

A second literature review was conducted to explore the importance of context-based design activity. As an outcome, design tools were proposed using participative design techniques involving games and role playing. Through a second series of five laboratory and field studies, the proposed design tools were developed and iteratively evaluated. It was demonstrated



that the design tools could affect interaction design activity, but further work is still required on improving one of the applicability criteria - 'organisational survival'.

These findings demonstrated that interaction designers can effectively produce their own design data using the design tools provided that this design activity is situated within the context of an interaction design problem. It has also been shown that if interaction design tools are to be effective they should satisfy all specificity and applicability criteria established in this inquiry.

Keywords:

Interaction design, Interface design tools, Interface design guidelines, Human-Computer Interaction, Consumer Product Technologies, Specificity and applicability criteria, Participative design.

To my Grandfather

Towards developing and improving effective interaction design tools

by

John V H Bonner

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of Doctor of  
Philosophy of Loughborough University

4 July 2002

© by John V H Bonner 2002

Contents

---

Volume 1

**1 Introduction..... 7**

1.1 Rationale for research..... 7

1.2 Funding and management ..... 9

1.3 Scope of research ..... 9

1.4 What is interaction design? ..... 12

1.5 Structure of the thesis..... 13

**2 A literature review on existing interface design guidelines ..... 16**

2.1 Introduction..... 16

2.2 Aims and objectives ..... 17

2.3 Establishing effectiveness criteria..... 17

2.4 Selection of interface design guidelines..... 20

2.5 Specificity of existing guidelines ..... 21

2.5.1 Content emphasis..... 22

2.5.2 Structure and presentation..... 25

2.5.3 Relevance of existing guidelines against ‘ergonomics criteria’..... 26

2.5.4 Relevance of existing guidelines to novel interactive devices..... 28

2.6 Applicability of existing guidelines ..... 34

2.6.1 Usability of interface guidelines..... 35

2.6.2 Reliability and validity ..... 37

2.6.3 Relevance to proposed target audience..... 38

2.6.4 Organisational survival ..... 41

2.7 Critical review ..... 43

2.8 Conclusions ..... 46

2.9 Methodology for more effective guidelines ..... 48

**3 Towards the development and use of CPT design guidelines ..... 51**

3.1 Introduction..... 51

3.2 Aims and objectives ..... 51

3.3 Study 1 - Development of CPT guidelines from the literature..... 51

3.3.1 Guideline development methodology..... 52

3.3.2 Critique of guidelines and development process ..... 54

3.4 Study 2 - Investigating applicability through interface design ..... 56

3.4.1 Devising a design research study..... 56

3.4.2 Method..... 60

3.4.3 Prototype solutions ..... 60

3.4.4 Results..... 64

3.4.5 Discussion ..... 70

3.5 Conclusions ..... 74



<b>4</b>	<b>Producing CPT design guidelines from usability evaluations.....</b>	<b>75</b>
4.1	Introduction.....	75
4.2	Aims and objectives .....	75
4.3	Study 3 - Evaluation of high level CPT prototypes.....	75
4.3.1	Method.....	76
4.3.2	Experimental design .....	78
4.3.3	Results.....	82
4.3.4	Guidelines derived from study .....	89
4.3.5	Discussion .....	93
4.4	Study 4 - Evaluation of low level CPT prototypes .....	95
4.4.1	Methodology and results .....	96
4.4.2	Final Prototype .....	100
4.4.3	Guidelines derived from study .....	101
4.4.4	Discussion .....	102
4.5	Conclusions .....	103
<b>5</b>	<b>Evaluating CPT guidelines in commercial design projects .....</b>	<b>105</b>
5.1	Introduction.....	105
5.2	Aims and objectives .....	105
5.3	Study 5 - Evaluating guidelines based on designers' requests .....	106
5.3.1	Method.....	106
5.3.2	Results.....	107
5.3.3	Discussion .....	111
5.4	Development of guideline design requirements .....	113
5.5	Study 6 - Evaluating the effectiveness of prototype CPT guidelines.....	119
5.5.1	Methodology and results .....	119
5.5.2	Discussion and consideration of more recent research .....	122
5.6	Conclusions .....	125
<b>6</b>	<b>Exploring alternatives to guidelines.....</b>	<b>127</b>
6.1	Introduction.....	127
6.2	Aims and objectives .....	127
6.3	Study 7 - The use of research-based design data in UK design practices .....	127
6.3.1	Method.....	128
6.3.2	Results.....	129
6.3.3	Discussion .....	134
6.4	Study 8 - Comparing two design cultures.....	136
6.4.1	Method.....	137
6.4.2	Electrolux findings .....	137
6.4.3	Danfoss findings.....	140
6.4.4	Comparing the two organisations.....	144
6.4.5	Potential effectiveness of participative design tools.....	145
6.5	Conclusions .....	148

<b>7</b>	<b>Exploring the theory and application of context in design .....</b>	<b>150</b>
7.1	Introduction.....	150
7.2	Aims and objectives .....	150
7.3	Theoretical and applied approaches to the study of context and design.....	151
7.3.1	Activity theory.....	151
7.3.2	Distributed cognition.....	155
7.3.3	Situated action model .....	156
7.3.4	Reflection-in-action .....	158
7.3.5	Using awareness of context in design.....	159
7.3.6	Contextual design.....	161
7.3.7	Scenario design.....	161
7.3.8	Interaction analysis .....	162
7.4	Comparing theoretical and applied approaches to design and context.....	163
7.5	Design tool framework .....	168
7.6	Conclusions .....	170
<b>8</b>	<b>Iterative development and evaluation of interaction design tools .....</b>	<b>171</b>
8.1	Introduction.....	171
8.2	Aims and objectives .....	171
8.3	Methodology .....	172
8.4	Rationale for design tools .....	173
8.4.1	Card sorting .....	174
8.4.2	Inspection based evaluation tool.....	177
8.4.3	Scenario design tool .....	180
8.5	Reconsidering the effectiveness criteria .....	180
8.6	Study 9 - Evaluation of the card sorting design tool.....	183
8.6.1	Evaluation plan.....	184
8.6.2	Observations .....	185
8.6.3	Reflections .....	192
8.7	Study 10 - Student user trials with design tools .....	193
8.7.1	Evaluation plan for MA students.....	194
8.7.2	Observations .....	196
8.7.3	Changes made to BA students' evaluation .....	199
8.7.4	BA student evaluation plan .....	199
8.7.5	Observations .....	200
8.7.6	Reflections .....	201
8.8	Study 11 - Perceived applicability of three inspection methods .....	202
8.8.1	Evaluation plan.....	203
8.8.2	Observations .....	208
8.8.3	Reflections .....	210
8.9	Conclusions .....	211
<b>9</b>	<b>Final evaluation of design tools.....</b>	<b>212</b>
9.1	Introduction.....	212
9.2	Aims and objectives .....	212
9.3	Design tool handbook .....	212
9.4	Study 12 - Evaluation of design tools in practice .....	215
9.4.1	Evaluation plan.....	215
9.4.2	Observations .....	220
9.4.3	Reflections .....	235
9.5	Study 13 - Investigating applicability factors further .....	237
9.5.1	Evaluation plan.....	237



9.5.2 Observations ..... 238

9.5.3 Reflections ..... 242

9.6 Conclusions ..... 243

**10 Conclusions ..... 245**

10.1 Introduction..... 245

10.2 Inquiry review..... 245

10.3 Recent research ..... 246

10.3.1 The use of guidelines and standards ..... 247

10.3.2 Use of design tools ..... 248

10.3.3 Bridging the gap using context..... 252

10.4 Critique of thesis :..... 254

10.4.1 Methodology ..... 254

10.4.2 Assessment criteria ..... 256

10.5 Contribution to knowledge ..... 258

10.6 Future directions..... 259

10.7 Final thoughts ..... 263

**11 References ..... 265**

**12 Acknowledgements..... 286**

Volume 2

**13 Appendices**

Overview ..... 1

13.1 Published Guidelines in ‘Human Factors in Consumer Products’ ..... 2

13.2 Interaction designer’s diary ..... 22

13.3 Investigative questions and evaluation methods for Study 3..... 32

13.4 Draft CPT guidelines..... 49

13.5 Design tools handbook..... 113

13.6 Internal Electrolux design report..... 269

Tables

---

Table 2.1 Guidelines effectiveness criteria ..... 20

Table 2.2 Sources of guidelines used in literature review ..... 21

Table 2.3 Summary of findings from the literature review ..... 47

Table 3.1 Guidelines found and unavailable during prototyping ..... 70

Table 4.1 Example question with corresponding capture and analytical methods ..... 77

Table 4.2 Experimental conditions ..... 78

Table 4.3 Recall and recognition tests for display objects ..... 79

Table 4.4 Best aspects of interface conditions ..... 84

Table 4.5 Worst aspects of interface conditions ..... 84

Table 4.6 Reported common mistakes for interface conditions ..... 85

Table 4.7 Reported impressions and suggested changes for interface conditions ..... 86

Table 4.8 Correlation in interaction level rating scores between  
novel and control conditions ..... 87

Table 4.9 Mean task completion times in minutes ..... 89

Table 4.10 Guidelines derived from audio display interface condition ..... 90

Table 4.11 Guidelines derived from drag and drop interface condition ..... 91

Table 4.12 Guidelines derived from animated object interface condition ..... 92

Table 4.13 Comparison of experimental design between Studies 3 and 4 ..... 95

Table 4.14 Ranking of selection feedback variants ..... 98

Table 4.15 Ranking of multiple selection variables ..... 99

Table 4.16 Ranking of animation icon formats ..... 100

Table 4.17 Ranking of dragging behaviour variants ..... 100

Table 4.18 Guidelines derived from quick and dirty studies ..... 102

Table 5.1 Summary of design recommendations for prototype CPT guidelines ..... 119

Table 6.1 Response frequencies for design specialisms ..... 129

Table 6.2 Type of interface design involvement ..... 130

Table 6.3 Design role in interface design ..... 131

Table 6.4 Response frequencies for gaining 'user information' ..... 132

Table 6.5 Response frequencies for user testing ..... 132

Table 6.6 Response frequencies for paper based ergonomics resources ..... 133

Table 6.7 Comparing MMI design tools against effectiveness criteria ..... 147

Table 8.1 Brief description of cognitive dimensions ..... 178

Table 8.2 Revised effectiveness criteria ..... 183

Table 8.3 Examples of cognitive dimensions structured around quantitative values.. 199

Table 9.1 Summary of suggestions to improve design tools ..... 242

Table 10.1 Critique of Säde (2001) study using effectiveness criteria ..... 249

Table 10.2 Critique of Andre *et al* (2001) study using effectiveness criteria ..... 251



## Figures

---

Figure 1.1 Framework illustrating scope of research .....	10
Figure 1.2 Flow diagram of research studies.....	14
Figure 2.1 Development and evaluation of CPT guidelines methodology .....	50
Figure 3.1 QOC Notation .....	59
Figure 3.2 Animated Object Display.....	61
Figure 3.3 Drag and Drop Touch Screen Display .....	62
Figure 3.4 Auditory based screen Display .....	63
Figure 3.5 Control Screen Display .....	63
Figure 3.6 QOC analysis of decision making path for icon dragging behaviour.....	65
Figure 3.7 Single option and criterion for problem solving.....	68
Figure 4.1 Icons selected through participant consultation.....	97
Figure 4.2 Prototype for 'adaptable' microwave interface .....	101
Figure 5.1 Sample of database-derived guidelines.....	108
Figure 5.2 Example of guidelines offered at design options .....	114
Figure 5.3 Example of guidelines offered as a checklist .....	115
Figure 5.4 Example of guidelines offered as discussion and comment .....	116
Figure 5.5 Example of a guideline providing evidence for the use of metaphors.....	117
Figure 5.6 Example of guideline rating scale.....	118
Figure 7.1 Basic activity structure (Kuutti 1996 p28) .....	152
Figure 7.2 Drag and Drop interaction style example .....	154
Figure 7.3 Animated Object interaction style example .....	155
Figure 7.4 New framework to gain design data and guidelines.....	169
Figure 8.1 Action research model for design tools evaluation.....	173
Figure 8.2 Example of original (top) and revised (bottom) cards.....	184
Figure 8.3 Example of student project.....	201
Figure 9.1 Procedural relationship of design tools .....	214
Figure 9.2 Participants involved in card sorting exercise .....	216
Figure 9.3 Tabs created by designers.....	218
Figure 9.4 Tab board in use.....	219
Figure 9.5 Tabs placed on prototype interface .....	220
Figure 9.6 Example of task map and function filter .....	228
Figure 10.1 Layered communication model .....	261

# 1 Introduction

---

*The source of invention and new design knowledge are not in the design of cookbooks and menus, but out in the vegetable patch. Bill Stumpf (1984, p35)*

Consumer product technologies (CPT) are becoming more 'interactive'. Computer-based technology is pervading conventional and new consumer products, often converging with telecommunications (Hood 1997, Poole and Simon 1997). Current examples of this type of product are mobile phones, personal digital assistants (PDA), interactive and web-enabled television. Future consumer products have also been suggested, such as smart domestic devices (Norman 1999) information appliances (Bergman 2000) and disposable micromachines that use cheap printing technology to provide throwaway microprocessors (Jacobson 2001). Whatever the future holds, designers of these emerging products need to account for the more abstract concerns of *interaction design*. Because of the potential novelty of these devices, the discipline of interaction design needs to develop its own declarative and procedural knowledge to facilitate the design of usable user interfaces and experiences for CPT.

## 1.1 Rationale for research

This research began by addressing the question,

*Can effective interface design guidelines be produced for use in the design of future consumer product technologies (CPT)?*

This question was important because criticism had already been levelled at a closely related discipline, Human-Computer Interaction (HCI). Brooks (1991) argued that HCI specialists provide deficient knowledge as it relates poorly to design decision making. Kuutti (1996, p 18) stated, 'researchers are unable to provide grounded advice on how such interfaces [direct manipulation]



should be developed'. Carroll (1991, p 1) commented, 'some of the most seminal and momentous user interface design work of the past 25 years made no explicit use of psychology at all'. Ironically, it has been suggested that theoretical HCI knowledge lags behind commercial practice knowledge by, in some cases, two decades (Sutcliffe and Carroll 1999). There exists a knowledge transfer 'gap' between HCI research and design practice, such as software engineering. This research aimed to prevent a similar situation developing, within the relatively young field of interaction design by reducing the gap between research domains such as human factors or ergonomics, human computer interaction, psychology and sociology and the practice disciplines of interaction design and industrial design. But how should this be achieved?

Gardiner and Christie (1987) proposed three approaches to reducing this gap:

- selecting an existing medium for transferring the knowledge
- identifying general research approaches appropriate for the development of future interfaces
- present research findings in suitable form, either as standards or in more creative, individual ways.

Instilled in these approaches is the assumption that existing theoretical cognitive psychology only need be appropriately packaged and focussed towards future design applications. This assumption is strongly contested in this thesis through two sets of short and progressive studies (Studies 1 - 6 and 9 - 13). Through the first set of studies it became clear that using design guidelines as a link or bridging mechanism between research and practice domains was not viable. This led to a shift to considering what should constitute an appropriate bridging mechanism. Studies 9 - 13 explored an alternative in the form of design tools. This shift required a different methodological approach. The first set of studies relied on classical experimental design and analysis. This gave way to action research methods

in the second set of studies, in which observation was combined with personal intervention. This was necessary as many of the studies were practice-based and, therefore, dictated by operational constraints within the field.

### **1.2 Funding and management**

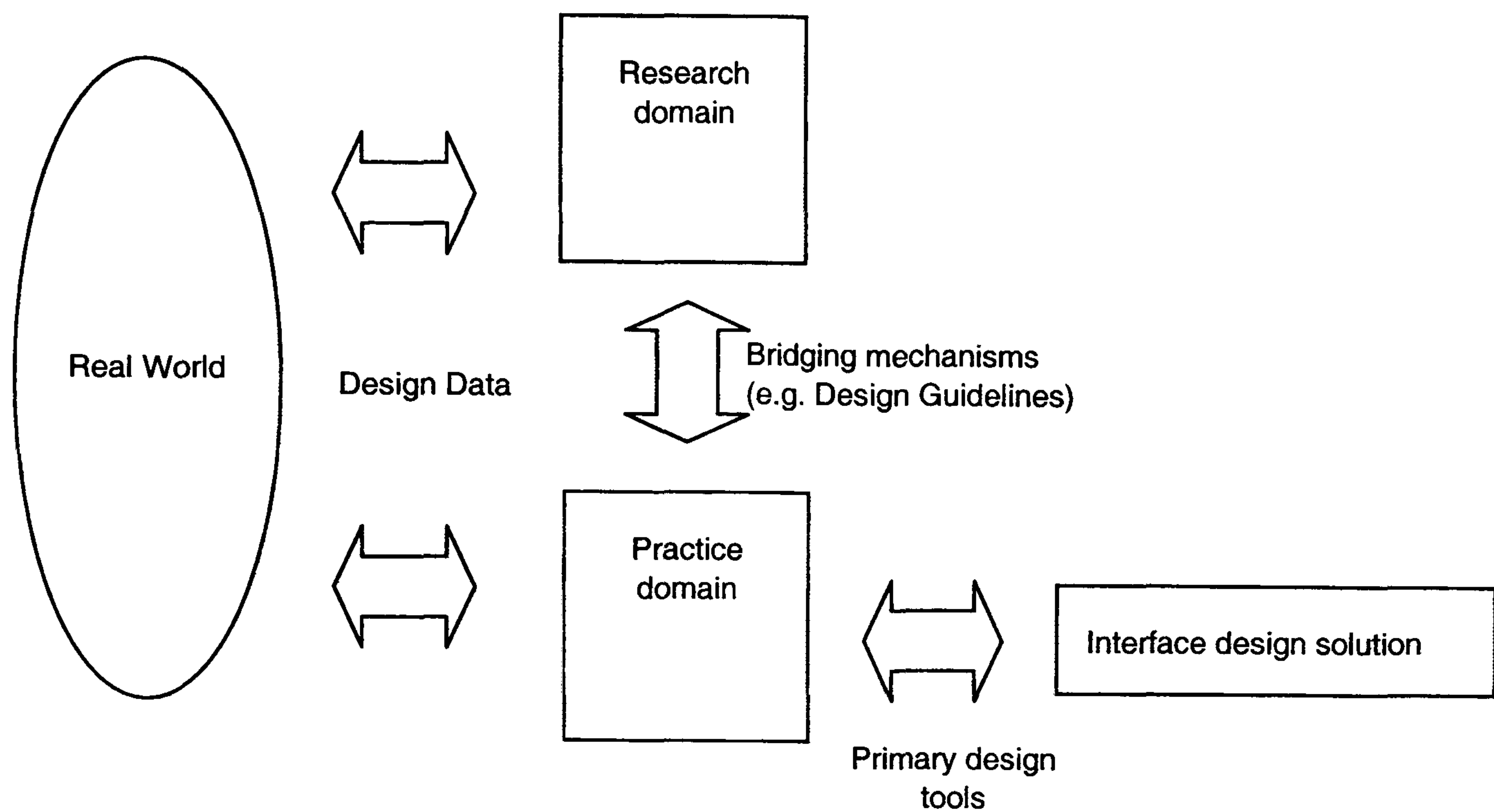
The development of the design tools was funded by an EPSRC research contract (GR/L 01787) with the author being the principal investigator. The project was assessed as having made a 'useful contribution to the field' (Alpha 2). Gaining commercial benefit was a condition for funding of this research from the EPSRC. Two organisations, Electrolux and Raychem, agreed to collaborate. Unfortunately, Raychem had to formally withdraw from the project due to the closure of the design group about 12 months into the programme, so other organisations including NCR and BT were introduced. Electrolux remained the key collaborating partner throughout the research investigation.

This research was carried out on a part-time basis. The empirical research was conducted over the first three years (1996 - 1999). The writing-up of the thesis was then carried out, after an 18 months gap, which was due to a move to a new teaching post in a different University.

### **1.3 Scope of research**

The research framework, presented in Figure 1.1, illustrates the key areas of concern and defines the boundaries and scope of this research. Certain terms used in this thesis have specific meaning. These terms are contained within Figure 1.1 and are defined below.





**Figure 1.1 Framework illustrating scope of research**

Within the research domain, disciplines such as HCI and ergonomics make abstractions of real world phenomena through observation, analysis and evaluation to form theoretical and applied prepositions and representations. Some of these abstractions can form what has been termed in this thesis 'design data'. This is any form of declarative or procedural knowledge that could influence or support interface design decision making, particularly issues related to usability. Examples of these data could be:

- findings from empirical studies; for example, the usability of deep and shallow menu structures
- findings from studies of different types of usability evaluation methods
- commonly held beliefs about design, designers and design practice.

This knowledge can be documented in research papers and texts but may also be retained as tacit knowledge by researchers and transferred between designers through collaborative working.

Different forms of bridging mechanisms can exist between research and practice domains to ensure that research-based design data are incorporated into interface design activity. Dix *et al* (1998) suggested there are three forms of interface design support: guidelines; analytic and predictive models of user behaviour; and evaluation techniques. Design guidelines are by far the most commonly produced and widely accepted type of bridging mechanism. These can be either expert-centred, through the use of design heuristics, or literature based through the use of prescriptive design rules or guidelines.

Usability practitioners often collaborate with designers by providing design data derived from the research domain by acting personally as the bridging mechanism and often play an important role in ensuring that research-based design data are used (Bias and Mayhew 1994; Trenner and Bawa 1998).

Whilst this is usually more effective than literature-based guidelines, due to the effect of personal immediacy, it can also be problematic as designers and researchers are culturally different. See, for example, studies identified in the literature review, Section 2.6.4 and observations made in Study 12, Section 9.3.2.

The scope of this research was limited to interface design activity carried out without the presence or support of usability experts where there *is a resulting dependency on literature-based bridging mechanisms*. Guidelines were initially selected as they offered the most flexible 'vehicle' to evaluate the effectiveness of a literature-based bridging mechanism. Other forms of literature-based bridging mechanisms exist, such as claims analysis, which provide options and justifications to designers (Sutcliffe 1999). These depend, however, on HCI-based usage scenarios and were therefore outside the scope of novel CPT applications. Interface design standards also exist for graphical user interfaces (GUI) but are currently not available for CPT.

Designers also gather design data from the real world. In commercial design teams this is likely to be in the form of marketing reports, product



development strategies, findings from focus group sessions, assumptions or preconceived ideas about end users, and cultural values held by the design group or the organisation. Through these sources designers use their own design data to support interface design decision making. It was not the intention to develop *primary* design tools, taken in this thesis to mean indigenous tools, such as drawing tools, CAD software and prototyping tools. For these types of tools, designers are expected to have competence and a deep understanding of their purpose.

The focus of this research was on the use and effectiveness of *secondary* design tools. Examples of these design tools are anthropometric tables and interface design guidelines. Secondary tools support design decision making but implicit or embedded in them is the recognition that design data are derived outside of the practice domain and typically accepted *prima facie*. Designers use them to make design decisions without necessarily knowing their derivation or theoretical background.

## 1.4 What is interaction design?

The young discipline of interaction design has a similar focus to that of HCI. Indeed, the term is gaining acceptance as a replacement for the discipline title of HCI (Barnard *et al* 2001; Preece *et al* 2002) as it is not restrained by the term 'computer'. Others view it as being a sub-discipline of product design (Säde, 2001; Weed 1996). Educational institutions that teach interaction design usually reflect a strong design orientation. Courses in interaction design differ from conventional HCI courses in curriculum content with a stronger emphasis on industrial and graphic design. The Interaction Design Institute, Ivrea, Italy, define interaction design as being, '... a new discipline: a fusion of aesthetics and culture, technology and the human sciences. It concerns the design both of the services these technologies might offer, and the quality of our experience of interacting with them (quote from Website).' Carnegie Mellon University, School of Design, USA, define it as concerning 'itself with the way people use and interact with technology: in software, appliances,

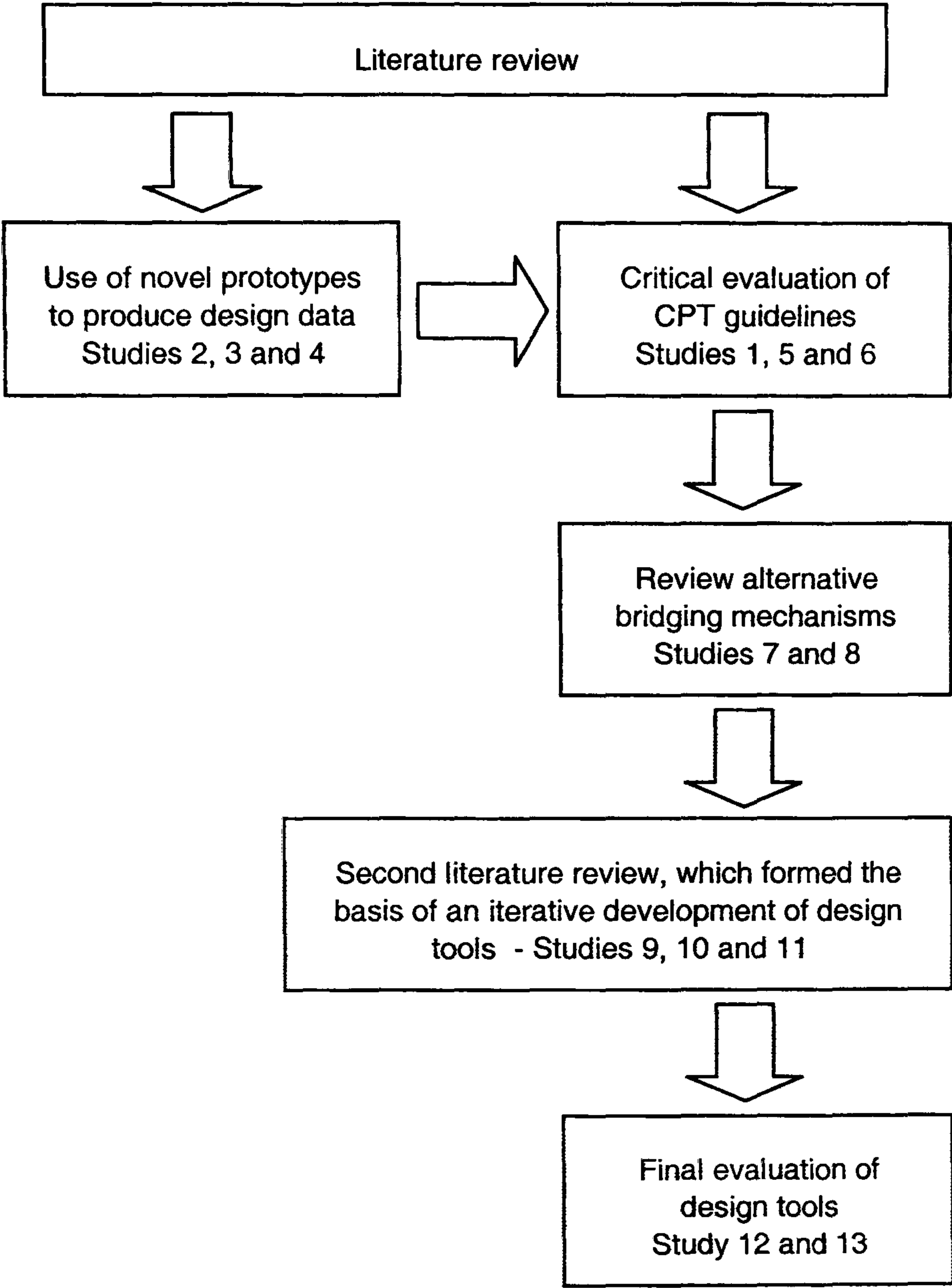
new media, and information spaces. We emphasize a collaborative design process which incorporates the skills of communication design, industrial design, cognitive psychology, and computer science, among others...(quote from Website) '. Few interaction design courses currently exist (Mountford and Johnson 1997), reflecting its fledgling status.

The emphasis of interaction design is on designing the quality of 'user experiences' (Alben 1996) by increasing the need to consider the product as an 'event' rather than an artefact. Equal value is placed on producing 'user experiences' through pure creativity as through experimental analysis. What separates interaction design from HCI is the emphasis on creativity and design. What separates it from industrial design is the concern with designing and evaluating abstract concepts of interactivity.

### 1.5 Structure of the thesis

A flow diagram illustrating the thirteen studies is presented in Figure 1.2 to assist navigation through the studies and to illustrate the relationship between them and to show how they cluster around common objectives. Most of the studies are presented in their chronological order, but Studies 1, 5 and 6 overlapped with Studies 2, 3 and 4.





**Figure 1.2 Flow diagram of research studies**

The research began with a literature review (Chapter 2) exploring the effectiveness of current interface design guidelines and establishing the effectiveness criteria for improving their usage. This was followed by Study 1, developing specific guidelines drawn from the literature that could be used in the design of CPT. Concurrently, Studies 2, 3 and 4 reported on the development and evaluation of novel prototypical interfaces, produced with the aim of developing specific guidelines for CPT. The guidelines produced through these studies were then evaluated in Studies 5 and 6. The accumulative evidence prompted the search for new bridging mechanisms in Findings from Studies 7 and 8, along with a more specific second literature review, provided design requirements from which a set of design tools were proposed.

The tools were iteratively refined through Studies 9, 10 and 11, and a full evaluation testing on a live design project was undertaken in Study 12, followed by a final, brief review (Study 13) exploring further how design tools might survive in an organisation.

The thesis concludes with a critical review (Chapter 10) of recent publications related to guidelines and similar design tools to those developed through this research. The contribution to knowledge is then presented, which states that firstly, despite improving the specificity and applicability of the CPT guidelines, passive, non-bespoke design guidelines have still been shown to have little impact on interaction design activity. Secondly, interaction designers can effectively produce their own design data using the design tools provided that this design activity is situated within the context of an interaction design problem. Finally, it has also been shown that if interaction design tools are to be effective they should satisfy all specificity and applicability criteria established in this inquiry.

The thesis concludes with suggestions for further research; interaction design research should test new design tools through in-the-field design activity rather than attempting to develop tools first and then attempt to apply them to the design process.

## 2 A literature review on existing interface design guidelines

---

### 2.1 Introduction

This literature review reports on what is known about existing interface design guidelines in terms of their use, scope and effectiveness. In order to address these issues, the review was carried out using a number of progressive stages.

The first stage was to understand what is meant by 'effective' guidelines and how this has been interpreted and defined in the literature. This helped derive operational effectiveness criteria by which firm critical judgements on existing interface design guidelines could be made.

The second stage identified existing interface design guidelines. Most of the reviewed guidelines were derived from the field of Human Computer Interaction (HCI). The guidelines were generally sought from popular or seminal reference texts that are readily available to software engineers, programmers and user-interface design experts. Other important or popular interface design guidelines covering 'hardware' based interfaces (physical control and display devices) were also sought as they have direct relevance to consumer product technology (CPT). The review also identified research using novel or 'cutting-edge' interactive technologies that could be applied to CPT in the future, even if the design data were not presented in the form of guidelines.

The third stage compared and critically assessed a selected set of interface guidelines and research reports against the 'effectiveness' criteria derived at the first stage. This helped to identify knowledge gaps and suggest where new design knowledge may be required. It also helped in understanding how future guidelines could be produced and implemented for interaction



designers. Finally, comparing existing guidelines and emerging interaction design knowledge against the criteria provided a rationale and justification for developing new interface design guidelines for CPT.

The findings from this literature review were critically assessed and some assumptions about future CPT guidelines were considered. This critique provided underpinning for a research methodology, proposed at the end of this chapter, to develop and evaluate interface design guidelines specifically tailored towards future CPT.

### 2.2 Aims and objectives

- To produce a set of 'effectiveness criteria', obtained from a synthesis of similar studies and reviews reported in the literature where guidelines were assessed
- To assess reports in the literature of current use and effectiveness of interface design guidelines
- To consider the relevance of existing guidelines for the design of future CPT
- Then, in light of these findings and if appropriate, to propose a methodology for the development of future suitable guidelines

### 2.3 Establishing effectiveness criteria

Carroll (1991, p2) states that for any applied science to be contextually relevant depends on two requirements: *specificity* where the content is relevant and *applicability* where the 'use of the science must conform to the processes of application in the target domain'. Effectiveness criteria were therefore selected by making clear distinctions between 'content' and criteria related to the successful implementation of any given content within a target domain. Specific, nested criteria were introduced under these two broad headings by reviewing studies that examined bridging mechanisms between research and practice in HCI.

First, in establishing more detailed 'specificity' criteria, descriptions of guideline types were sought. Henninger *et al* (1995) divided guidelines between technology-centric

---

Button [Keyboard] combinations are tricky to use, and should only be used when no other method can be found, or the function is a rarely used one. In preference use function keys or 'stacked control keys'. (Martin and Eastman 1996 p35)

---

and abstract or general-purpose guidelines.

---

Define and use terms consistently.  
Once a term is defined, especially when it may have a common English meaning which is not intended, use the term consistently in the defined sense. (Brown 1988 p59)

---

Tetzlaff and Schwartz (1991) made more detailed distinctions, suggesting different approaches: general principles such as usability heuristics; specific interface design detail; guidelines to produce a consistent 'house-style'; guidelines on interface presentation and behaviour; while others address performance characteristics. Therefore guidelines offer, either explicitly or implicitly, an emphasis or accent in terms of their intended use of application. The effectiveness or suitability can be considered in terms of their *content emphasis*. Linked to this is how the content is presented. The literature review did not reveal any direct comparative studies between different sets of guidelines. Nevertheless, *content structure and presentation* was regarded as an important effectiveness measure of guideline sets and was included as a criterion.

A common criticism of guidelines is their lack of *validity* (interpretation of the meaning and significance of a guideline) *and reliability* (consistency in applying guidelines to different design problems). Bastien and Scapin (1995) argue that inspection methods (the use of HCI experts to evaluate the usability of an interface without user involvement) which included the use of guidelines need to be critically assessed in terms of validity, thoroughness and reliability. They propose a set of empirically developed 'ergonomics



criteria' which provides a framework for the accurate description and classification of likely usability problems. Through a series of studies they have developed, refined and validated the effectiveness of their ergonomics criteria to be unambiguous and comprehensive. They argue that any form of inspection method can only be effective if all 'ergonomics criteria' are considered during the evaluation process. Therefore *ergonomics criteria* (described in detail in Section 2.5.3) provided an opportunity to evaluate the effectiveness of guidelines in terms of their relevance or coverage in identifying usability problems.

Critical judgement was also required on how far existing guidelines would be *relevant to novel CPT*, in terms of content. To assess this, the scope and relevance of existing guidelines needs to be judged against the likely design data needed for emerging CPT.

Very few studies were found that addressed broad aspects of the applicability of interface design guidelines. Mosier and Smith (1986) addressed applicability in a narrow sense. They reviewed the *usability* of guidelines by examining how users (Smith and Mosier 1984) rated them by answering a questionnaire. Usability was defined in terms of readability of guidelines; finding specific information and applying the guidelines; and the perceived usefulness of guidelines

Grudin and Poltrock (1989) and Grudin (1991) have observed how interface design decisions are made across large organisations where effective multi-disciplinary teams are critical if good user interfaces are to be developed. These studies suggest that conflicting demands between disciplines often exist, where compromises and trade-offs form an important part of the decision making process. This suggests criteria for measuring applicability must include *relevance to the intended target audience* and, in a broader perspective, must consider how guidelines might *survive* within a design organisation. No studies were found that specifically attempted to

understand external or organisational factors that would affect the acceptance or use of guidelines. The environment in which guidelines have to compete is complex and constantly changing and little is known about the real impact that interface design guidelines have on improving usability. Table 2.1 summarises these criteria under the two generic headings of specificity and applicability.

Specificity Criteria	Applicability Criteria
Content emphasis	Usability of guidelines
Structure and presentation	Reliability and validity of guidelines
Relevance of guidelines against 'ergonomics criteria'	Relevance to proposed target audience
Relevance of guidelines to novel CPT	Likelihood of survival or usage in organisational or design context

Table 2.1 Guidelines effectiveness criteria

2.4 Selection of interface design guidelines

A comprehensive collection of *generic* interface design guidelines was drawn from the HCI and ergonomics literature in order to assess their specificity (obviously applicability could not be assessed without conducting empirical studies). Generic guidelines were defined as interface design guidance not specific to any particular application, usage or appliance. 'Application-specific' guidelines (for example telephone-based design guidelines) or 'user-specific' guidelines (such as users with special needs) were found but not included in the review. The reason for this was that in these cases the content was too focussed, making any meaningful comparative specificity judgements difficult.

Fifteen texts, providing interface design guidelines, were reviewed. The texts selected were either purchasable or accessible through libraries to designers (the term 'designer' was used in a broad sense and includes those attempting to seek any form of good hardware or software interface design data). Less



accessible texts were also selected if they were professionally well regarded and cited. The texts used in this review are listed in Table 2.2 along with a brief descriptive summary. This list is not exhaustive and there are many other important and seminal guidelines but their inclusion would have been unlikely to add any further significant findings.

HCI Guidelines	Ergonomics 'hardware-based' Guidelines
Brown (1989), Human-Computer Interface Design, <i>Provides general interface concepts, design of interface components and addresses some dialogue issues such as error handling.</i>	Clark and Corlett (1984) The Ergonomics of Workplaces and Machines: A Design Manual. <i>Guidelines can be accessed through decision tree flow charts.</i>
Flower and Stanwick (1995) The GUI Style Guide, <i>Detailed coverage of interface design mainly at the device level.</i>	Galer 1987, Applied Ergonomics Handbook, <i>Discusses the physical design and use of displays and control and basic principles for panel and machine layout.</i>
Galitz (1993), The Essential Guide to User Interface Design, <i>Mixes process and implementation issues with practical interface design guidelines at device level.</i>	Helander 1995 A Guide to the Ergonomics of Manufacturing, <i>Design guidelines on control, symbols, labels and visual displays provided.</i>
Martin and Eastman, (1996) The User Interface Design Book, <i>Places strong emphasis on the creative interface design process. Less emphasis on the design of interface components and greater consideration of multi-media aspects such as graphics and sound.</i>	Sanders and McCormick 1993 (7 <sup>th</sup> Edition) Human Factors in Engineering and Design, <i>Extensive, often empirically based design advice provided within around 'input' and 'output' processes between the user and machine.</i>
Ravden and Johnson (1989), Evaluating the Usability of Human Computer Interfaces, <i>Provides an evaluation checklist methodology based on nine interface design principles.</i>	Woodson <i>et al</i> 1992 Human Factors Design Handbook: Information and Guidelines for the design of systems, facilities, equipment and products for human use, <i>Provides a strongly diagrammatic data source covering applied systems design, environmental design and physiological response to the environment. Some conventional display and control design guidelines provided.</i>
Rivlen, Lewis and Cooper (1990), Guidelines for Screen Design, <i>Guidelines focus on screen layout, text and graphic with additional consideration of navigational and control issues.</i>	
Shneiderman (1992), Designing the User Interface, A <i>comprehensive coverage of research related to the use and design of Graphical User Interfaces. Design 'guidance' is presented through findings of reported studies.</i>	
Smith and Mosier (1984). Guidelines for Designing User Interface Software, <i>One of the earliest examples of interface guidelines, provides a comprehensive set of guidelines that were published before the wide adoption of Graphical User Interfaces.</i>	

Table 2.2 Sources of guidelines used in literature review

2.5 Specificity of existing guidelines

The intended audience for all of these texts was not necessarily the same. Some guidelines provide supplementary evidence or knowledge to support



scholarly learning, while others were designed to support an interface design process. This was a consideration when drawing comparisons from the review.

### 2.5.1 **Content emphasis**

In examining the first effectiveness criterion, *content emphasis*, the most obvious distinction between the texts was the division between guidelines intended for hardware-based and software-based interface design and Table 2.2 reflects this. The HCI texts focus heavily on screen design, although some consideration was given to hardware devices such as the keyboard and mouse. In contrast the ergonomics based guidelines discuss physical control and display devices.

HCI design guidelines were more prevalent than hardware design guidelines. The design and manufacture of hardware devices requires high tooling and manufacturing costs and design decisions are made infrequently resulting in slow and evolutionary changes. Thus interaction protocols and principles have matured and stabilised over many years and are based predominantly on physical rather than abstract interaction design problems.

Conversely, the use of software-based interfaces is relatively young. Control and display devices are more diverse, complex and abstract. Interface design skills are required across a wider spectrum of software design activity and therefore non-HCI specialists make interface design decisions. Software interface designers continually change and evolve screen designs and interaction styles with relative ease, resulting in a proliferation of interface design solutions. This creates a greater demand for software based interface design guidelines.

This contrast between stable, intuitive, slow-evolving hardware design guidelines and abstract and fast-evolving software design guidelines is reflected in the specificity of the texts. The ergonomics-based texts (Clark and

Corlett, 1984; Galer 1987; Helander 1995; Sanders and McCormick 1992; and Woodson 1992) are generally consistent in their emphasis and coverage. All texts divide between human output (control) and human input (display) with a strong emphasis on the physical form, behaviour and appearance of such devices. Only Sanders and McCormick (1992) address the design issues of computer based devices such as keyboards and touch screens as well as considering, albeit briefly, more novel interaction devices, such as speech-activated controls. However, few design issues are considered.

Most of the design guidelines are 'engineering' based with data on mechanical displays using moving pointers and scales, the use and design of control devices such as knobs, dials,

---

Dials are satisfactory for most information except warnings, and complex and stored information. Circular dials are better for comparison and rate of change.  
(Clark and Corlett 1984 p66)

---

cranks, levers and pedals revealing that most of the empirical work is related to post-second world war industrial machinery. While these data are still relevant, software-based control and display devices are superseding them.

The HCI guideline texts (Brown 1989; Flower and Stanwick 1995; Galitz, 1993; Martin and Eastman, 1996; Ravden and Johnson 1989; Rivlen, Lewis and Cooper 1990; Shneiderman 1992; and Smith and Mosier 1986) are more variable in content and emphasis. There were obvious marked contrasts between early guidelines (e.g. Smith and Mosier, 1986) which offer very prescriptive guidelines against a backdrop of specific software applications such as data entry, and the more recent design 'guidance' texts such as Martin and Eastman (1996). In the latter, the process of interface design is framed within a creative design process in which the designer has more screen design freedom. More recent HCI texts also reflect current technology trends, in particular multi-media (Flower and Stanwick 1995 and Martin and



Eastman, 1996) illustrating how ephemeral software design guidelines can be.

---

A better interface might use a slider where the level could be set by dragging a pointer up and down a scale. This can be taken a stage further. The application could display a picture of a light bulb in a panel, at five different stages of illumination. Then, as the slider was moved around, the bulb would glow brighter or go dimmer. This approach has the advantage of both communicating exactly what the control does and making the interface with the application more attractive to use. (Martin and Eastman 1996 p76)

---

Many of the texts constructed their information around the component elements of a screen, the selection and design of elements such as menus, icons, pointers and error messages, for example, (Brown 1989; Flower and Stanwick 1995; Galitz, 1993; Rivlen, Lewis and Cooper 1990; Shneiderman 1992; and Smith and Mosier 1986).

- 
- Consideration in Screen Design
    - Where to place information
    - What information to place on a screen
    - How to place information on the screen
      - Upper and mixed case fonts
      - Special symbols
      - Field Captions/Data fields
  - Data Entry Screens
  - Inquiry Screens
  - Multipurpose Screens
  - Question and Answer Screens
  - Menu Screens

[Sample chapter headings (Galitz 1993)]

---

Only Ravden and Johnson (1989) attempt any radical departure from this approach in raising design issues through the use of nine key design principles (headings are provided in the example below) and presenting them in the form of checklists.

- 
- Visual clarity
  - Consistency
  - Compatibility
  - Informative Feedback
  - Explicitness
  - Appropriate Functionality
  - Flexibility and Control
  - Error prevention and correction
  - User guidance and support

(Ravden and Johnson 1989)

---



This text, however, emphasises an evaluation rather than a design approach. Galitz (1993) mixes content emphasis between good interface design principles based on graphic design and cognitive psychology principles such as functional grouping of text or graphic elements, and around task activity such as information entry and modification and data entry.

HCI guidelines are derived from psychological knowledge such as that found in theories of information processing, from empirical evidence, the author's personal experiences, or from an amalgam of different published guideline sources. Most texts do not present an explicit rationale for their guidelines, but are usually supported with citations or a brief explanation from the findings of a study. Rivlen, Lewis and Cooper (1990) were the exception in providing a rationale alongside a related set of guidelines. Most texts are intended to be read by practitioners seeking design or evaluation advice, although the content emphasis with Shneiderman (1992) is to provide more theoretical and empirical evidence of human-computer interaction as a research discipline, intended to provide a deeper understanding of the underlying principles of HCI.

### **2.5.2 Structure and presentation**

The structure and presentation of the texts vary enormously and are usually related to their intended function. Some texts (Galer 1987; Helander 1995; Sanders and McCormick 1992) are not necessarily designed as reference texts and offer design data more as background support, whereas, Clark and Corlett (1984) and Woodson (1992) are clearly intended to be read as design data reference texts. This is reflected in the presentation style of the latter examples, with more dependence upon examples and graphical images. Clark and Corlett (1984) use decision tree structures to organise their data, providing access to data using a problem-solving approach. The provision of good design examples is also more common in texts intended to support the interface designer or practitioner.

Some texts have been intentionally written to support design data searching, allowing the user to seek out specific guidelines or to support a systematic process of considering design guidelines. Brown (1989) structures the content around 'checklist' statements, presenting these in bold followed by a supporting 'one-liner' guideline, which is then discussed in more detail using one or two paragraphs. At the end of the text, examples are provided to articulate design thinking.

---

### 7.11 Input Field Definition

*Show the length of entry fields*

Use input prompts to indicate clearly how many input characters are required. Some methods for indicating field lengths are solid-line underscores....

#### **Examples:**

In all these examples, the field length indicators are replaced by input characters as data are entered into the field.

#### **Broken-line underscores**

(preferred) ACCOUNT NUMBER: \_ \_ \_ \_

(Brown 1989 p126)

---

Ravden and Johnson (1989) provide three case studies in order to illustrate how this checklist methodology can be used. Rivlen, Lewis and Cooper (1990) used the left-hand page for guidelines and the right for a supporting rationale.

Other texts map the design process and use this as a structuring mechanism. Galitz uses a 12-step design process (Step 1 - Know your User or Client; Step 2 Understand the business function...) a mechanism also used by Martin and Eastman, (1996) but in a less structured manner. Clark and Corlett (1984) introduce new topics with a flow chart; for example, criteria for control design. Typically, though, chapter and section headings in most texts relate to device or component interface characteristics.

### **2.5.3 Relevance of existing guidelines against 'ergonomics criteria'**

Scapin and Bastien (1997) and Bastien and Scapin (1995a, 1995b) developed the notion of 'ergonomics criteria' through a series of empirical studies. They



argued that the metrics on which the guidelines are based, are variable in terms of their comprehensibility, independence between guidelines and reliability. This was the only study found that attempted to produce a set of empirically based ergonomics criteria for evaluating the 'ergonomic quality' of interactive systems. This was achieved by producing a set of complete and valid ergonomics criteria. A large set of recommendations and production rules was produced and then, through a series of iterative user trials, they compiled a commonly agreed set of mutually exclusive criteria.

Once the criteria were formed, reliability was tested by asking human factors specialists and non-specialists to match known usability problems against the criteria. This allowed distinctions between the criteria to be improved through the refinement of definitions and the inclusion of new examples. The effectiveness of these improved criteria was then tested by comparing the number of usability problems identified with and without the support of the criteria. The use of the criteria resulted in significantly more problems being identified. Finally, they studied their effectiveness against the design principles used in ISO 9241 - Part 10, and found that using the ergonomics criteria uncovered more usability problems. From these studies, they produced eight main criteria, sub-divided into eighteen elementary criteria. The main headings for the criteria were:

- Guidance - related to the ability to advise, instruct and guide the user through the interaction process
- Workload - interface elements that reduce perceptual or cognitive load
- Explicit Control - processing and control of explicit user actions
- Adaptability - system ability to behave contextually and according to user needs
- Error Management - reducing errors and error recovery
- Consistency - consistent use of codes, formats etc. in different contexts



- Significance of codes - relationships between codes and their meaning
- Compatibility - task and user match

The authors argued that the criteria developed are complete (i.e., covering current guidelines) and distinct (i.e., independent of one another). They stated that these criteria are only applicable to ergonomics criteria of known interactive systems and may not be relevant for new interaction technologies. The criteria were used to 'benchmark' the coverage of different guideline sets. In using these criteria, it was possible to judge the guidelines sets on aspects of completeness and independence between guideline sets.

In terms of comparing the ergonomics criteria against the selected texts, only one was found to be almost identical (Ravden and Johnston, 1989). The checklist criteria have some semantic differences in terms and emphasis in scope. For example Ravden and Johnston use a criterion known as visual clarity which cuts across 'guidance' and 'significance of codes' in the Scapin and Bastien criteria. The other texts generally considered a subset of the Scapin and Bastien criteria. Some texts mixed device based guidelines with 'ergonomics criteria' based guidelines; Brown (1989), for example, discusses the use of control and display devices in one chapter while in another covers the design of error messages and error correction.

The texts that were most removed from the ergonomics criteria were those that focus on screen design (Flower and Stanwick 1995; Galitz, 1993; Rivlen, Lewis and Cooper 1990). These texts placed less emphasis on criteria such as workload, adaptability and error management. Screen design issues were considered predominantly within the 'guidance' criterion.

### **2.5.4 Relevance of existing guidelines to novel interactive devices**

Existing guidelines, particularly the majority of the ergonomics guidelines reviewed above, discuss control and display devices that are now rarely found on CPT. This claim could also be made with HCI guidelines, which

assume interaction is conducted with a mouse, keyboard and large screen. Many of the reviewed guidelines could have been dismissed as out-of-date or irrelevant and thus ineffective. However, without some clear predictions of where future interactive technology may lead, it is unwise to dismiss device level (software or hardware) guidelines out of hand. Gardiner and Christie (1987) recognised that future trends in technology will create knowledge gaps where new user interface design principles will be required. The identification of these knowledge gaps, they argue, should trigger further demand for pure and applied research in cognitive psychology. Some predictions or technology trends were presented in order to suggest where these gaps might occur. This allowed suggestions to be made as to what type of theoretical and empirical knowledge would be required in order that new guidelines could be derived. Therefore, by using technology trends as a predictor of interface design knowledge needs, the disparity between existing guidelines and the need for future interface design guidelines could be determined.

This approach to using technology trends to identify guidelines was used in this literature review. To determine whether existing guidelines contain relevant design data for novel interactive devices, they were assessed against the literature discussing emerging interactive technologies that has also highlighted design data needs.

A literature review was, therefore, carried out to identify common technology trends. Key journals and conference proceedings, such as Computer Human Interaction (CHI) conferences, which consider this type of research activity, were selected. In identifying examples of novel interactive technology, the review was limited to recent conference proceedings to reflect state-of-the-art research. In this way, it was possible to draw some high level indicators as to where design data may be required and to make some comparative judgements against existing guidelines. From this review three broad trends emerged:



- The use of intelligent or 'smart' interactive technology in consumer products
- The demand for more 'natural' interaction styles
- Enquiry into the use of novel input and output modalities

These papers were not being judged on the quality of their research rigour or even their viability. What was important, was to gather design experiences, expressed in the form of a need for, or expression of, design knowledge, during the design and development of novel interactive devices.

### *Intelligent or 'smart' interactive technology*

Much of the theoretical research in this area stems from the field of Artificial Intelligence (AI). This field has largely fallen from favour because of the narrowness of tangible applications. Attention has more recently shifted to the concept of intelligent agents. Research by Maes (1995) explored the use of autonomous agents, described as computational systems that live in a complex system. These agents can sense and act autonomously in their environment and, by doing so, can achieve certain goals. In one particular project users can interact with animated characters, such as a dog, by interacting with the virtual environment. Specific research initiatives in the UK are beginning to be realised. A recently funded research project by the Design Council at Edinburgh University, 'Increasing Information Intensity: Towards "Intelligent" Products', has begun to explore the impact of intelligence within the consumer product domain and has proposed ways in which intelligence might be used in different consumer product applications. (Fleck *et al*, 1997).

Product intelligence does not have to be manifested in software. According to Black and Buur (1996) intelligence can be embedded within a solid user interface (SUI). They discuss research developing intelligent polymers that



form and deform depending upon the product-state. Studies related to the design and usability of 'smart' products are still elusive but two studies on the implications of using intelligence in interactive products were found. Normura and Brownlow (1996) reported on how intelligence was used in the development of the Sharp Logicook microwave oven and on the design implications of designing and evaluating an intelligent interactive product. They discussed some of the design problems faced by the product designer, in particular, being able to model all possible interaction scenarios. Burford and Baber (1994) studied the introduction of an adaptive interface for an auto-teller machine and reported on user acceptance. They found that if the product was to be successful it was important that users were made explicitly aware that they were using an adaptive interface. These studies suggest that the introduction of intelligent product interfaces will need to be evolutionary and incremental. Interaction designers will need to know what type of intelligence representation will be required to ensure reliable and valid user acceptance and task match.

### *The use of 'natural' interaction styles*

Natural interaction styles purport to offer usability through the exploitation of more 'conventional' human-to-human interaction. In normal social interaction we use a complex set of utterances, gestures and expressions to convey meaning. Researchers have attempted to carry-over these social communication codes to human-computer interaction in an effort to make interaction more natural and intuitive. A large number of human-computer interaction styles have been developed which mimic 'natural' modalities such as verbal communication.

One interaction style that has received considerable attention is the use of auditory icons. These are everyday sounds that are used to convey information and provide feedback about a computer's actions (Gaver 1989). The use of voice-controlled interfaces or automatic speech recognition is beginning to emerge, particularly with phone-based interfaces. Halstead-

## 2 - Literature review on existing guidelines

Nussloch (1989) presented a set of interface design guidelines offering support on when to use phone-based interfaces and how different interaction styles should be considered, such as the use of verbal prompts. The guidelines offer interaction style choices, such as types of navigation control, but offer little in terms of comparative or qualitative judgement. Where this is offered, the design advice stems from basic interface design principles such as matching controls to frequency of use - 'the designer should order the options for choice according to their use frequency'. Research by Sidhu and Coyle (1995) reported a similar set of design principles such as good feedback, reduced complexity and easy to understand prompts.

Schumacher *et al* (1995) argued that phone-based interfaces suffer from poor usability because of a lack of design guidelines and they provided a set of comprehensive guidelines to help improve future interfaces. Guidelines are grouped around user activity such as navigation and the use of feedback. The authors do briefly discuss the limitations of guidelines and how adherence to them does not guarantee a usable interface. They also admit that the guidelines have not been empirically validated and argue that, despite this, such design knowledge is required to reduce the design cycle time.

### *Novel input and output modalities*

Using hand gestures to communicate with computers has received some degree of research attention. Hand gestures take advantage of such features as naturalness, adaptability and dexterity; users therefore think more about directly affecting a task. Gesture control and interaction can be static or dynamic in terms of being an interaction dialogue. Dynamic gestures are powerful in that they allow humans to combine a number of poses and easily communicate complex input messages. For example, it is possible to specify an object, the operation to be performed on the object and additional parameters by means of one gesture (Bolt 1992, Lu *et al* 1995, Shimada 1995). These studies emphasised the need for the user to understand the



'grammatical' structure of gestures and the importance of gaining helpful feedback; otherwise users can not estimate whether their input has been realised by the interface.

Some work has been undertaken in using field-sensing devices where a room can be designed with monitoring equipment that tracks the location of its occupants. This can have important implications for using the body as a control device (Zimmerman 1995). Research by Maes (1995) used a full body interface in which the user stands in front of a large screen and a video camera picks up body form and replicates an image of the user within a virtual world. She found that, despite the 'naturalness' of body movement within the virtual world, users still required human intervention in the form of tips and hints to understand some of the behavioural models within the world. She also found that representational fidelity of virtual objects was not critical but the behavioural actions of the objects were. For example when interacting with a virtual dog, realism was only achieved if the user felt the dog was exhibiting natural behaviour. Eye movement was particularly important in conveying this.

Attempts have also been made to develop interactive communication devices for children, in which designers (Piernot et 1995) attempted to blur the boundaries between software and hardware elements. For example, when a communication card is inserted into the device a virtual representation of the card appears on the screen. They suggest that novel interactive devices with very specific tasks will demand specific types of user centred design skills, particularly as children cannot articulate their needs in the same way as adults.

In a similar project, Fitzmaurice *et al* (1995) developed the concept of a graspable user interface. This allowed direct control of virtual objects through physical handles, a small brick. The brick or bricks allow the user to control virtual objects by physical manipulation and selection of different

bricks. Using this principle, prototypes for drawing tools have been developed. The authors present many issues inherent in type of interaction style that need to be addressed, for example, how bricks should communicate with each other and how functional attributes should be assigned between virtual and real objects.

This review of potential CPT provided an opportunity to identify where knowledge gaps may occur. The use of 'intelligence' in products, however broadly or narrowly this is defined, does appear to be imminent. The studies previously identified suggest that design guidelines will be required to help the selection or development of effective modelling methods and techniques to design and evaluate future interaction styles using computer-based intelligence. Guidelines may also be required that address way in which users perceive and use intelligent products. These may also extend to elaborating on the boundaries and limitations of 'intelligence' that should be used in an intelligent product and how it should be articulated to the user (Bonner 1999). The common challenge with 'natural' interaction styles is to provide appropriate and adequate feedback. Although the interaction style concept may be natural, interaction protocols still have to be learnt.

Designers will need to be provided with information on the boundaries and limitations that users are prepared to accept or will tolerate with these new forms of interaction styles. The use of novel input and output modalities is difficult to predict but evidence from a reasonably mature interaction modality, such as audio displays, suggests that very little is still known or generally available to the designer.

### **2.6 Applicability of existing guidelines**

The literature review revealed very few guideline-specific studies addressing applicability. Those that were found divided between appraisals of design guidelines themselves in terms of their own 'usability', or comparative studies of guidelines against other forms of interface design methods.



### 2.6.1 Usability of interface guidelines

Mosier and Smith (1986) conducted a questionnaire survey of 130 respondents who had used the Smith and Aucella (1983, in Mosier and Smith 1986) design guidelines. About 50% of the respondents were human factors specialists and 25% were either systems analysts or software designers. They found that different professional groups used the guidelines in different ways. Managers and software designers tended to skim the report, while human factors specialists read it in more depth. Software designers found the guidelines to be less helpful, although this was based on a sample of ten respondents. The guidelines were used in more cases as a method to establish user requirements or as an aid during the design process than as an evaluation tool. Problems identified relating to the use of the guidelines included:

- respondents mentioned a general reluctance by other members of their design group to use human factors expertise
- the guidelines were too general to be applied
- lack of guidelines for new technologies
- inability to locate relevant guidelines
- guidelines were not set in any context in terms of their value or importance and users would have liked assistance in choosing guidelines for use in design.

The authors suggested that these problems could be overcome by the development of software support tools to assist in providing access to and tailoring of guidelines. This approach was adopted by Henninger *et al* (1995) who developed a guideline database that allows personal contextual knowledge to be attached to guidelines to increase effectiveness and enable reuse in subsequent design projects.

In a study by Tetzlaff and Schwartz (1991) nine participants were asked to design an screen-based interface using paper and pencil while using a set of

interface design guidelines. This study is interesting in that the design process and the support of that process were important components of the research. The participants had to develop interface solutions to a range of scenarios. They observed that the guidelines were very poor at supporting these processes and provided little integrative information. The guidelines were most effective at the detailed interface design stage. They also noted that the participants adopted two different strategies in using the guidelines. A small group did read the introductory material before embarking on the guidelines whilst the majority read until they found what they needed to know. They found that pictorial examples were heavily relied upon, often at the expense of the accompanying text and, similar to the De Souza, *et al* (1990) study, the designers extrapolated beyond the original intentions of the guidelines. Some of the participants did report that the experiment 'artificially induced guideline usage'. Nonetheless, the guidelines were more effective than the authors originally anticipated and the designers found them valuable in providing examples of conforming interface design and in providing explicit strategies for tackling interface design. They concluded however, that guidelines should only be used as part of a development toolkit and that other methods of evaluation should also take place.

Other studies were found in which comparisons were made between guidelines and other forms of design or evaluation advice. Jeffries *et al* (1991) assessed four methods of user interface evaluation: heuristic evaluations, cognitive walkthroughs, software guidelines and usability evaluations. They found heuristic evaluation (evaluating from an HCI expert point of view) identified the highest number of interface problems, including the most serious. Whereas guidelines were good at identifying recurring and general problems, they also state that guidelines, no matter how specific, could not have identified certain problems, as they were highly specific to the application. When designers are faced with a range of guidelines, they may only refer to topic areas they believe to be applicable and not consider other relevant information. Cuomo and Bowen (1994) in a similar study,



comparing three 'evaluation' techniques (software guidelines, cognitive walkthrough and heuristic evaluation) made more subtle comparisons between evaluation methods. In their study they explored usability problems against 'stages of user activity' during human-computer interaction such as the user's intended action, the execution of that act and interpretation of the act. They found that, although guidelines found the widest range of usability problems, they were the least effective in terms of identifying problems that had a noticeable effect on user performance.

### **2.6.2 Reliability and validity**

De Souza, *et al* (1990), recognising there had been little research in validating the effectiveness of guidelines, observed three interface designers designing a menu system using guidelines and found a range of error types was made. In broad terms, these were either misinterpretations of the guidelines, or correct interpretations incorrectly implemented. They also found that designers were not passive users of guidelines; rather they actively used their own knowledge and experience in conjunction with the guidelines. Although they found guidelines were useful and contributed towards improved usability, they suggested that a means of expressing the validity of a guideline, both in terms of its scientific basis and also in the design consequences of using the information, may improve its correct selection and implementation. Unfortunately, this was not expanded upon.

Lavery *et al* (1997) argued that the validity of evaluation methods needs to be defined in terms of their 'quality' (most accurately identifying the number of usability problems later found in user testing). They state that many evaluation methods have been compared using usability predictions or the number of usability problems revealed. They argue that validity combined with quality is a more effective criterion because it is more important to identify and predict usability problems that will cause problems later than it is to reveal all usability problems without any form of qualitative measure. This suggests that the effectiveness of guidelines also needs to consider their

ability to consider or anticipate usability problems that have a real long-term impact on usability, even when guidelines can claim to be comprehensive, valid and reliable.

Finally the literature review did not reveal studies that attempted to measure the reliability of guidelines - that is how effective guidelines were at being used in different problem solving situations.

### **2.6.3 *Relevance to proposed target audience***

For guidelines to act as a bridging mechanism between research and practice there has to be a degree of reciprocal empathy. The providers of design data (ergonomists and interaction design researchers) and the users of design data (for example interaction designers) have historically tackled problem solving differently. Bishop and Guinness (1966 p 281) identified many overlapping domains of interest between the two disciplines but also identified differences in approach. They state, 'The designer approaches a problem in a systematic way, but is not structured by the scientific rigor of the human factors specialist. The industrial designer is more widely eclectic. His interest in a design problem and its final solution need not be as narrowly defined as that of the human factors specialist. He relates solutions to the complete resolution of the design problem, using widely diverse criteria.' They suggest, through examples, that these differences in approach can complement each other; the human factors specialist can offer problem convergence by defining or focusing the problem, in contrast to the industrial designer who offers problem divergence through a range of possible solutions.

Clearly, working methods differ. Industrial designers usually reason inductively, adopting a 'solution-focused' strategy with an initial idea created early on that guides the development process (Lawson 1980). While human factors specialists adopt a deductive reasoning approach. Wilson and Corlett (1990, p2) state 'Ergonomics is both a science and a technology and thus has a



need of techniques for both data collection ... and application'. Both disciplines place a high importance on the 'end user', although this is less rigorous in industrial design where Grant and Fox (1992) suggest that there is a strong reliance on an aesthetic criterion for design evaluation and designers use only themselves to evaluate their work.

Rouse *et al* (1991) explored how different disciplines designing complex flight-crew systems represented phenomena, problems and solutions. They considered the way in which different disciplines viewed the nature of phenomena and how areas of concern were framed. They divided phenomena between internal phenomena (which occur within the human such as decision making, social interaction and emotion) and external phenomena (such as natural physical laws of nature such as the elasticity of metals). These phenomena could be framed within real or model worlds. Model worlds are abstractions from the real world and again disciplines can be distinguished in terms of how they use real and model world data. They suggested that engineers have an external, real world perspective, whereas computer scientists use an external, model world, as programs or indeed the interface, are functional abstracted models of the real world. Human factors specialists, they suggested, have an internal, real world perspective. They have a focus on human activity within an applied real world context and therefore depend on empirical evidence and experimental models.

If we assume that industrial designers reflect engineering philosophy, we can see that the cultural differences may centre on the consideration of internal and external phenomenon. Whilst these are sweeping generalisations, they help may to draw distinctions. Distinctions between design and ergonomics do appear to be less clear cut, but there is some evidence that designers are more willing to understand and use human factors methods and techniques rather than the reverse. Fulton (1993) provided evidence of this stating that within their organisation, industrial designers are departing from intuitive approaches by introducing simple user trials to challenge their instinctive

assumptions about a product. McClelland (1990) suggested that there is a wide range of ergonomics methods and techniques that can be adopted by industrial designers but argued that their acceptance will only occur when ergonomists integrate themselves and become part of the design, rather than the evaluation, process.

Boff (1988) argued that this cultural gap between the users and producers of design data is due to the designers' overall perception of costs and benefits. He argued that most new design activity is dependent on either adapting or using variants of previous design solutions (referred to as baselining). Designers will not expend energy on gaining any additional information beyond what they view as relevant. Therefore, for human factors data to be used, it must be integrated into these baseline solutions. He suggested that increasing the use of human factors data can only be achieved by raising the perceived value while reducing learning costs. He suggested that this can be resolved by an increase in more applied research and changing attitudes towards human factors.

The literature revealed little evidence or examples of practical, bridging mechanisms, which successfully fulfil the criteria of usability, reliability, survival and relevance. Designers, as predicted in this thesis, will increasingly become users of interface design data. But the evidence from the literature suggests that interface designers and researchers in interaction design both produce and use design data very differently.

There remains a requirement to understand the type of bridging mechanisms required, to ensure that interaction designers use research-based design data. Suggestions from the literature suggest that this 'bridging' could be achieved through a number of delivery mechanisms such as the integration of designers and human factors specialists (Haubner 1990). Gillan and Bias (1992) suggested that the content, structure and style of human factors data should be offered through educational tools. Gould *et al* (1991) accepted that



software engineers do not use user-centred methods, in spite of showing enthusiasm for them, and suggest that one of the ways of overcoming this is by setting long term usability targets for future application developments. While Boff (1988) suggested that human factors research should be demand driven by the users of the data, but also suggested that educating the users of this knowledge will still be required. Rouse *et al* (1991) made some specific suggestions for design support in flight-crew systems design. They suggested software support through awareness and tutoring functions along with technical information. However, none of these studies provided proven practical solutions that have been implemented and shown to be appropriate for designers. Guidelines have many shortcomings but, at this stage, still remain a useful 'vehicle' to understanding many of the deficiencies that still exist in bridging mechanisms between the two domains of application and research.

### 2.6.4 **Organisational survival**

Studies were sought that explored the designers' perception of usefulness with different design and evaluation methods. Klein and Brezovic (1986) explored the types of 'human performance data' needed by 50 systems designers to resolve interaction design problems based on their recollections of previous design projects. The findings suggested that designers rated literature based research findings poorly and favoured personal experience and prototype development. They comment, 'designers appear to avoid formal decision making...they are less interested in adding to their option set than in working with the best option of the ones they have identified...' (p 775). Hammond *et al*, (1983) in interviewing a small number of software designers, found that they were not familiar with human factors literature and also reported concern over the relevance of the information provided by in-house human factors specialists. They found that design guidance, either verbal or documented, was too low level or general to be of use. What they were seeking, and often did not get from the human factors specialists, was applied and contextually-relevant information. Much of the designers' user

requirements knowledge was gained from simplistic psychological models of user behaviour.

In a broader perspective, no studies were found that specifically explored the perceived usefulness of interface design guidelines within an organisational context, though one study was found that examined how interface design knowledge and expertise is managed within a large organisation. Grudin and Poltrock (1989) reviewed how different professional groups (software, industrial design and human factors engineers, technical writers, training developers and marketing representatives), with a vested interest in interface design, perceive and work with each other. The study revealed that there was a high degree of need for human factors information across a wide range of project specialisms (such as marketing and software engineering). Interestingly, a high incidence of external consulting rather than internal advice was reported to obtain the requisite information. The study did, however, report that strong relationships usually exist between human factors specialists and industrial design engineers, with the latter reporting a high level of human factors competence. Gould *et al* (1991) offered six reasons why software engineers do not take usability 'seriously'. They suggested that reward structures and working methods did not allow the serious consideration of usability, that usability can not be objectively measured and suggested that better tools for iterative design were required.

Grudin (1991) provided a detailed account of the organisational and design factors that inhibit the adoption of user-centred practices. He illustrated that, during product definition, other considerations such as product marketing will override any consideration of user requirements. He suggested that there are many organisational obstacles to identifying and obtaining feedback from users. The goals and motives that thrive within a software development project affect the implementation of usability goals. For example, goals to develop software code that are modular, reliable and



maintainable will be an important driver in dictating how an interface is designed.

### 2.7 Critical review

Evidence of good bridging mechanisms between research and practice was rare. Many studies revealed that human factors design data was not received favourably or easily adopted within design practice despite some consensus of its importance. Examples of effective guidelines were found. Alben *et al* (1994), and Teasley and Scholtz (1997) reported on the successful implementation of GUI standards and guidelines. It was assumed, based on this literature review, that their success could be attributed to having a known target audience and implementing the guidelines within a specific problem solving context. Non-bespoke guidelines appeared not to have been developed using user-centred methods. They were not specifically tailored towards the user requirements of a heterogeneous user population. Two texts, however, had introduced a user-centred design *approach* (Galitz 1993 and Shneiderman 1992) as an integral component of the guidelines but only in an adjunct fashion. Considering the user tended to be emphasised in the introductory part of the texts supported with suggestions on the importance of user evaluation in the conclusions. Users were rarely considered as an integrative part of the guidelines themselves.

The *content emphasis* of most guidelines was centred on the 'device' aspects of interface design (the design of menus and command lines for example) rather than using a designer-centred approach (for example providing guidelines as options with advantages and disadvantages). Results from this literature review suggest that if device-based guidelines continue to be published, they will become increasingly irrelevant for future CPT unless they consider the potentially divergent nature of new types of control and display devices that are beginning to emerge. Existing guidelines were found to be generally 'prescriptively' focussed, providing design advice with little discussion about the potential conflicts or trade-offs that may need to be considered, or how

interface design should be considered within the context of task activity or differing user needs. Providing guidelines that have a more designer-centred approach could overcome these difficulties.

A variety of formats was found in *structuring and presenting* guidelines. It would appear that some form of rationale is important to support any given design guideline and help reduce possible misinterpretations or allow the designer to make more informed judgements about the selection or priority of using a particular guideline. Beyond this, the review did not suggest effective structuring or presentation styles.

Studies that assessed applicability tended to be narrow laboratory controlled evaluation studies (De Souza, *et al* 1990 and Tetzlaff and Schwartz 1991) where the *usability of guidelines* were assessed against contrived design problems. This form of experimental design makes their findings difficult to extrapolate to real world design problems that designers face. For example, some texts offered design advice (see example below).

---

If a number of items in the text form a list then you can show this visually by placing them in a column (Illustration xx)

Differentiate lists from the surrounding text by spatial positioning such indentation (same Illustration xx)

Do not use different colours for things which are the same, for example, alternating line in different colours

(Rivlen, Lewis and Davies-Cooper 1990 p74)

---

However, it is unlikely that it would be used as design advice in the real world. A designer would not read guidelines, like the example given, and then decide to use this formatting structure while also adhering to the advice. The more likely scenario is that the guidelines would be read retrospectively, once a table had been implemented; thus they support evaluation rather than creative design activity. Field-based studies by Boff (1988) Grudin (1991) Klein and Brezovic and Hammond (1983), suggested that designers were naturally resistant to adding new design problems and prefer to use their



own experiential knowledge above imported *secondary* knowledge. This suggests guidelines will only be effective if they evolve and become integrative and solution focused towards problem solving. To overcome these problems, guidelines need to be carefully constructed around designer-centred activities. Brooks (1991, p50) argues that HCI specialists will only be effective if they can provide 'high level analyses useful for evaluating the impact of major design decisions' and 'information that suggests actual design rather than simply general design guidelines or metrics for evaluation'.

Evidence was also found that industrial designers were prepared to move into human factors territory. There appeared to be a dichotomy between the need for human factors knowledge by designers but a general scepticism about the support provided by human factors specialists themselves. Some unproven solutions were offered on how this problem could be overcome. These ranged from quite radical proposals such as organisational restructuring to providing tailorable and adaptable design data

Scapin and Bastien (1997) derived their ergonomic criteria from design rules and then evaluated them as an evaluation tool, not as a tool to support creative design activity. Their studies have helped to improve the *validity* of usability evaluation methods, but ergonomics criteria do not assist as a design tool. However, evidence from the literature suggests that designers are unlikely to frame design problems using abstract criteria such as 'error management' although these criteria have been proven to be useful during the evaluation. Research is required that identifies design as well as evaluation-based ergonomics criteria. This may be an important contributing factor to improving the effectiveness of guidelines. Studies such as Mosier and Smith (1986), De Souza *et al* (1990) and Tetzlaff and Schwartz (1991) found guidelines to have limited effectiveness, were open to misinterpretation or provided design data that were difficult to apply. This suggests that existing guidelines, which have been intended to support

design activity, are only realistically used to evaluate design solutions and do not provide creative solutions.

### 2.8 Conclusions

Effectiveness criteria were derived from the literature allowing guidelines to be evaluated using two criteria sets of *specificity* and *applicability*. Overall the current use, scope and effectiveness of existing guidelines was limited and these limitations are summarised in Table 2.3. Guidelines suffered from poor specificity because designers find them difficult to apply to their design problems. Evidence was found that the applicability of guidelines has little proven success beyond single, bespoke organisational needs. Furthermore, guidelines appear not to compete well against other design methods.

In re-visiting the original research question - *Can effective interface design guidelines be produced for use in the design of future consumer product technologies (CPT)?* - evidence from this literature review clearly suggested that effectiveness was limited. Design guidelines, however, remain an important repository for design data generated from the research domain. Therefore it could be argued that, while this is the case, improving their specificity and applicability remains an important objective.



Specificity Criteria	Applicability Criteria
<p><i>Content emphasis</i> - Most software texts divide between 'design process', 'device based' or 'interaction dialogue model' approaches for content emphasis. Clear differentiation between hardware and software interface guidelines was found. Existing device-based guidelines are becoming increasingly irrelevant for CPT.</p>	<p><i>Usability of guidelines</i> - Usability can be poor because guidelines are difficult to locate due to their 'generality'. Suggestion in the literature that adding personal and contextual knowledge to guidelines may help in improving usability (and re-use) of guidelines. Guidelines need to be written so that they can be used in a flexible manner.</p>
<p><i>Structure and presentation</i> - Great diversity in style and approach. Identifying an effective structure and presentation style for CPT guidelines still needs to be proven</p>	<p><i>Reliability and validity of guidelines</i> - Studies have explored the validity of guidelines but this measure alone may not improve effectiveness. Measures are required that identify usability problems that have the highest impact. Guidelines appear to be good at revealing specific types of usability problems.</p>
<p><i>Scope of usability issues identified</i>- 'Ergonomics criteria' based guidelines may have proven validity but still need to be made more appropriate and relevant as part of the design process as well as the evaluation process. Criteria may not extend to CPT interfaces</p>	<p><i>Relevance to proposed target audience</i> - Cultural differences found between designers and human factors disciplines. Conflict found between designers needing human factors data and being sceptical of human factors specialists to provide it. Designers more keen to move into human factors 'territory' than vice versa. No user-centred approaches found to designing guidelines</p>
<p><i>Support for novel interaction styles</i> - New interactive technologies will require new device-based guidelines, but also new knowledge on the users social and cognitive acceptance of new interaction styles.</p>	<p><i>Likelihood of survival or usage in organisational context</i> - Other design and evaluation methods will often be chosen in favour of guidelines. Relevance to target design problem often seen as a difficulty. Generally guidelines have a low survival rate.</p>

Table 2.3 Summary of findings from the literature review

Gaps in the literature were found that suggested where further research could take place. Guidelines were not specifically tailored towards CPT. The review of emerging technologies concluded that new interaction design data was required. Little is known about how users will perceive, interact with and use these new technologies and, indeed, what methods and techniques are best suited to gaining this knowledge. Guidelines have not been produced using user-centred design methods, which may contribute towards their ineffectiveness. Some evidence was found suggesting that the applicability of guidelines could be improved if guidelines were either integrated into, or supported, the designer’s experiential learning process. These gaps provided strong enough evidence that improvements could be made to the effectiveness of interface guidelines by assessing them against

the specificity and applicability criteria. However, the overall limitations of guidelines could not be ignored and therefore careful consideration was given to progressing this inquiry by using guidelines as a vehicle towards understanding how bridging mechanisms between research and practice domains could be improved.

In light of this, the research question was revised to:

*How can effective interface design data be produced for use in the design of future consumer product technologies (CPT)?*

### 2.9 Methodology for more effective guidelines

In order to answer this question, CPT design data needed to be developed and evaluated. One of the reasons for using guidelines was that Electrolux and Raychem were willing to collaborate on the condition that they gained design guidelines to support their future design development work in CPT. The proposed methodology provided four different strategies to developing and evaluating CPT guidelines while accommodating deficiencies in specificity and applicability identified in the literature. An account of this development process was published in Bonner (1997). By evaluating design guidelines in this way, it was hoped that an understanding about providing effective design data would be gained.

The first strategy (Study 1) was methodologically 'conventional' by deriving guidelines from other literature sources having direct relevance to CPT. To allow a more dynamic approach to developing guidelines, a database was used as an interim repository for 'raw design data' rather than collating design guidelines direct from the literature. The database was used as an interrogation tool allowing similar and related topic areas to be compared before producing design guidelines.



The second strategy (Study 2) involved gathering designer-centred requirements for new guidelines to improve applicability. This was achieved by analysing the needs of an interaction design while developing interface design prototypes of novel CPT, to identify what type of appropriate and relevant guidelines can adequately support design activity. In this way, it was possible to understand applicability issues more thoroughly.

The third strategy (Studies 3 and 4) investigated how effective it would be to proactively produce guidelines from design data derived directly from empirical usability evaluation studies specific to CPT. Empirical experiments were designed to gain specific design data suited to the development of new guidelines.

The fourth strategy (Studies 5 and 6) was to evaluate the CPT guidelines. Again a designer-centred approach was adopted to ensure that applicability as well as specificity criteria were considered. A number of product designers were interviewed to establish user requirements for interface design guidelines and to understand the organisational context in which guidelines potentially need to survive. To further support the development of user-centred guidelines, designers were asked to provide requests for the type of design guidelines they might require during an active or future interface design project. Guidelines were produced from designers' requests using design data from the database. Content was therefore driven by user requirements. This approach was quite a radical departure from conventional guideline evaluation methodologies found in the literature.

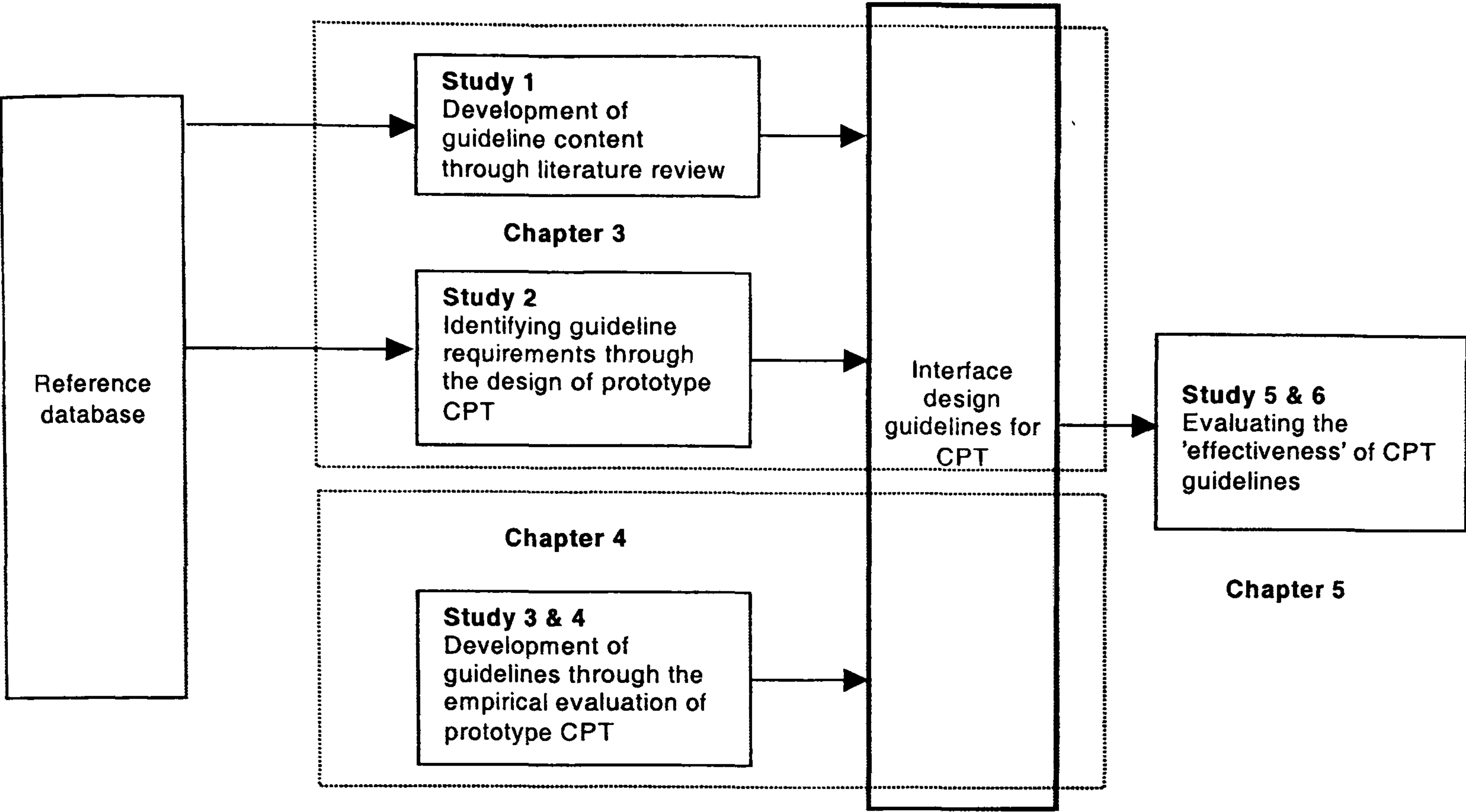


Figure 2.1 Development and evaluation of CPT guidelines methodology

Figure 2.1 provides a schematic representation of the methodology. Chapter 3 presents the first two studies, Study 1 - developing CPT guidelines from the literature and Study 2 - identifying designer requirements for CPT guidelines. Chapter 4 describes Studies 3 and 4 where evaluation trials were carried out to establish if specific CPT guidelines could be produced proactively. Chapter 5 presents Studies 5 and 6 where the CPT guidelines were evaluated. These studies formed four different strategies to developing CPT guidelines (represented by the bold box).



## 3 Towards the development and use of CPT design guidelines

---

### 3.1 Introduction

To understand more deeply how specificity and applicability factors might affect the effectiveness of new guidelines for consumer product technology (CPT), two studies were carried out, a development study and an observational study. Thus forming two of the four development strategies proposed in Section 2.9.

The first study reports on the development of specific CPT guidelines synthesised from the literature. The guidelines developed from this methodology are presented in Appendix 13.1 and were published in a book chapter entitled '*Towards Interface Design Guidelines for Consumer Products*' (Bonner 1998). A critique on the development process follows. The second study explored applicability issues, by taking a deeper look at real design activities and how design guidelines interact with this process.

### 3.2 Aims and objectives

- To gather and synthesise interface design data specific and relevant to CPT and begin to develop these data into design guidelines
- To critique the effectiveness of these new CPT guidelines and the data gathering process in light of the specificity criteria derived in Section 2.3
- To explore what factors affect the applicability of guidelines

### 3.3 Study 1 - Development of CPT guidelines from the literature

The objective of this study was to gather appropriate guidelines relevant to CPT, which could then be used for subsequent evaluation studies (Studies 5 and 6). The intention was to understand more precisely how guidelines should be developed and conveyed to interaction designers.

### 3.3.1 Guideline development methodology

The review strategy involved searches in three key areas:

- ergonomics and HCI texts providing key *design principles* that naturally extend towards CPT
- published *research studies* in the fields of HCI, ergonomics and industrial design concerned with conventional control and display devices that could be applied to CPT
- a review of the '*grey*' *literature* where the design and use of consumer products is discussed on less rigorous academic levels, and included non-refereed conference proceedings

Abstracts from papers and summaries were entered into a database (Filemaker-Pro). Reference papers were not classified using a content-based taxonomy as, at this stage, it was difficult to determine how these data might be used. Flexibility was key in terms of adding to and searching the database. For this reason, a simple classification system was adopted to enable searches on author, date of publication, paper title, and abstract text string matches. Searching on text strings provided a very flexible keyword search, allowing new reference material to be continually added to the database without the need to create new classifications. The database contained HCI and ergonomics interface design guidelines that were relevant or specific to CPT. Abstracts were entered into the database and over 600 reference titles were collected covering a period of approximately 10 years (1986 - 96).

Once the database had been populated, guidelines were then synthesised from relevant and appropriate papers using keyword searches. In light of the findings from the literature review in Chapter 2, the following rules for guideline development were adhered to:



- Design data were made as relevant to consumer products as possible, however at this stage this was limited to producing guidelines for conventional rather than emerging consumer product technologies to simplify the collating task.
- Attempts were made to phrase guidelines around designer-centred issues rather than ergonomic or usability criteria to increase their relevance to designers. This approach was taken in light of the perceived weakness of current guidelines as being relevant to design activity.
- As far as possible, guidelines were produced so they were 'product independent' to increase their relevance to a wider readership.

As already noted in the previous chapter, no real insights for good guideline structure and presentation had been gained. This meant that a degree of intuition was used to select an appropriate structure and presentation style. It was decided that guidelines should be as short and simple as possible. Sixteen 'headline' guidelines (italicised statements in Appendix 13.1) were produced to address broad usability issues allowing the reader to skim the key guidelines. Each guideline was supported with a rationale to allow more informed judgements to be made. References to other publications for further reading, were also included, allowing the reader to source more detailed material.

Guidelines were clustered around the following headings:

- *Appropriate contextual fit for the interface.* Guidelines discussed the importance of ensuring that product interfaces are suitable for the relevant user population and accommodate their initial and dynamically changing requirements.
- *Appropriate display devices and feedback.* These guidelines concentrated on ensuring consumer products provide relevant information on the user's actions, product state changes, guidance and support.

- *Appropriate control of the product- state and functions.* These guidelines focused on ensuring users are provided with adequate control mechanisms in order to carry out tasks effectively
- *Adopting a user centred design approach.* This section briefly discussed this design philosophy and directed the reader to more detailed publications for further reading.

#### **3.3.2 Critique of guidelines and development process**

The development of the design data database proved to be a useful learning exercise and resource. On reflection, data gathering lacked focus. It was difficult to identify what design data would eventually convert well into design guidelines. This led to uncertainty about implementing a robust taxonomy. Different taxonomies could have been adopted such as relevance to application domains, interaction models, ergonomics criteria or stages of the design process. However, adopting the wrong taxonomy would have resulted in data access problems or enforced predefined data structures within the guidelines. Although the use of text string searches resolved the problem, papers selected for the database were reviewed less critically as it was not as important to evaluate their relevance against any classification system.

The structure and presentation of the guidelines were not considered a high priority at this stage. Constraints in publishing the guidelines in a book meant that more 'adventurous' structures and presentation styles could not be explored. This problem was not thought to be important at this stage as it was hoped that a good structure would become clearer later when designers were involved in the guideline development process (see Chapter 5). The key objective at this stage was to generate design guidelines for further refinement. For this reason, this issue was deferred.

It was considered important that the content emphasis should be around concepts or processes that would be useful and understandable to interaction



designers. The 'ergonomics criteria' were heavily biased towards an evaluation context and therefore not thought appropriate if the guidelines were to be grounded in design rather than evaluation activity. In order to provide a set of comprehensive and simple design guidelines, relevant to a wide range of consumer products, each guideline became phrased more as a 'checklist' design statement rather than be related to design activity. This forced the guidelines to 'drift' towards general principles rather than be specific to design issues. While producing the guidelines, they became generalised in tone with much of the specific design guidance provided in the rationale, thus paradoxically reducing accessibility. This resulted in the guidelines remaining evaluation rather than design-centred guidelines

With each guideline a rationale for their inclusion was provided. Within this there was a danger of over simplifying design guidance, for example, in providing advice on adopting a user centred design approach. This was, partially, resolved by providing references for further reading but this inhibited accessibility to further design guidelines. The guidelines often could not be read at face value, as further reading was required to obtain more relevant information. Each rationale, as far as possible, provided examples of how such design guidelines could be implemented within the context of consumer products. Providing examples based purely on consumer product interfaces was difficult, as this type of design data was limited. Examples were often mixed between consumer and computer-based products where computer based interface design guidance were used to fill in the gaps. The obvious danger here was assuming that HCI design guidelines were transferable.

To conclude, while specific CPT design data were found, the author found it difficult to synthesise these data into guidelines. They could not be generalised beyond the specific domain from which they been had derived. Design guidance had to be phrased as broad 'commonly-known' design principles. Many trade-offs had to be made. Improving the relevance of the

guidelines was difficult, increasing their relevance across many products or application domains also increased their generality. Although, what must be taken into consideration is that the primary objective of this study was to produce content to allow subsequent improvements to be made through iterative evaluation. Nevertheless, the most difficult task was keeping the guidelines grounded within tangible and practical design-centred evidence or examples.

### 3.4 Study 2 - Investigating applicability through interface design

The literature review (Chapter 2) revealed that studies *had* evaluated the effectiveness of guidelines but only within narrowly constrained, experimentally based design problems. To redress this and understand how *all* applicability factors may affect the use of guidelines, a longitudinal interface design project was undertaken using a more natural, longer-term design task involving the development of novel interaction styles for future CPT. During this process, interface design decisions could be monitored and assessed against applicability criteria.

Prototype development, in this study, was therefore different to conventional prototype development outcomes. Developing prototypes to evaluate interactivity is common practice in HCI and human factors and advocated as an integral and important part of the interface design process (Gould and Lewis, 1985; Hix and Hartson, 1993; Newman and Lamming, 1995).

However, in this study, prototyping was used as a *vehicle* to produce new guidelines specific to CPT.

#### 3.4.1 Devising a design research study

The first task was to identify a potential CPT candidate for prototyping so that a range of novel interaction styles could be developed. A collaborative venture (through the EPSRC research contract) was set up with Electrolux Industrial Design Group in the UK who were willing to put development resources into exploring alternative interaction styles within their white



goods products. They had identified a washing machine panel (AEG Lavamat 6955) that consumers were finding difficult to use. Electrolux were starting to think strategically about superseding many of their conventional physical control and display devices with touch screen technology but were undecided about the functional form this technology should take.

This provided a research platform to develop speculative interaction styles while also beginning to understand how design guidelines could be provided for an organisation such as Electrolux. It was decided that any novel CPT interface proposed should have sufficient 'generic' qualities to be applied to other future Electrolux products. It was also agreed that prototypes should be based on available technology and that any proposed novel interfaces be evaluated against an existing washing machine interface, ensuring a more comprehensive comparative usability analysis between existing and novel interaction styles. Tentative user requirements for the novel CPT prototype interfaces were derived from a range of marketing reports. Electrolux had conducted extensive consumer surveys providing some understanding of what consumers wanted from future washing machines. Four internal reports were reviewed.

The first report investigated consumer perceptions towards technology. Generally respondents were wary of technology and did not like many of the innovative features presented to them. In the second report, a total of 200 consumers were asked to comment on a range of new appliances. The results indicated that out of the 45 positive statements collated from the interviews, the highest number of responses was attributed to statements such as "looks easy/simple to use" and "less busy/fewer controls". The third report examined the '50+' age range market to establish what the negative and positive interface characteristics would be for a particular concept machine. Some of the general conclusions were that interfaces were complicated and difficult to read; colours used on the display were attractive but not functional and lacked contrast and buttons were too small. The fourth report

presented the findings from a workshop comprising of three groups each with 3 single males and 6 females. The majority of respondents wanted machines that were simple to use and should not be like video recorders. However, a minority looked for products that required 'technological know-how'. Many respondents felt there were too many programmes available and only used 2-4 of them. Washing symbols were generally perceived to be easy to use.

The surveys lacked experimental rigour, with many of the reports being vague about the exact methods adopted, particularly the comparative surveys. However despite this, there appeared to be a level of consensus concerning user requirements. Users had a general distrust of technology and needed to feel in control with perceived 'ease of use' being important. Usability was a very strong selection criterion. There was a small minority of users who welcomed more control of their appliances.

Based on these findings from the survey reports and requirements, three variant interface prototypes (plus the existing washing machine panel) were proposed for development. These would use a touch screen, high-resolution display screen technology and exploit auditory displays. Concept design proposals were developed with the Electrolux UK design team before embarking on prototype development.

Electrolux agreed to fund a junior interaction designer to develop the prototypes who was based at Teesside University for six months. The findings from the development process and the usability evaluations were reported to Electrolux at the end of the study. The intention was to formally record design decisions during prototype development. This proved difficult to implement as the designer found it difficult to keep a record of his thoughts. Eventually, the designer agreed to keep a day-to-day diary of events but this quickly fell into disuse. The designer, however, compiled a



recollection of key design activities after the prototypes had been completed and this document was subsequently analysed.

Various analytical methods for the diary were considered such as IBIS (Issue-based information system) which has been extended to support graphical representation of a design space (Conklin and Yakemovic 1991). These methods, however, should be used during design activity for recording historically accurate records of decision making and therefore could not be used as a post-hoc tool. A more suitable and simple tool, focussing on the conclusions of design activity, which did support post-hoc analysis, was QOC (Questions, Options and Criteria; Maclean and Mckerlie, 1995). This method allows explicit representation of alternative design options and develops a rationale for the selection of one of these options based on criteria, see Figure 3.1. Design questions posed during the design process are examined in terms of the options (in effect the possible answers that could solve the question) and their corresponding criteria or reasons that argue for or against the possible options.

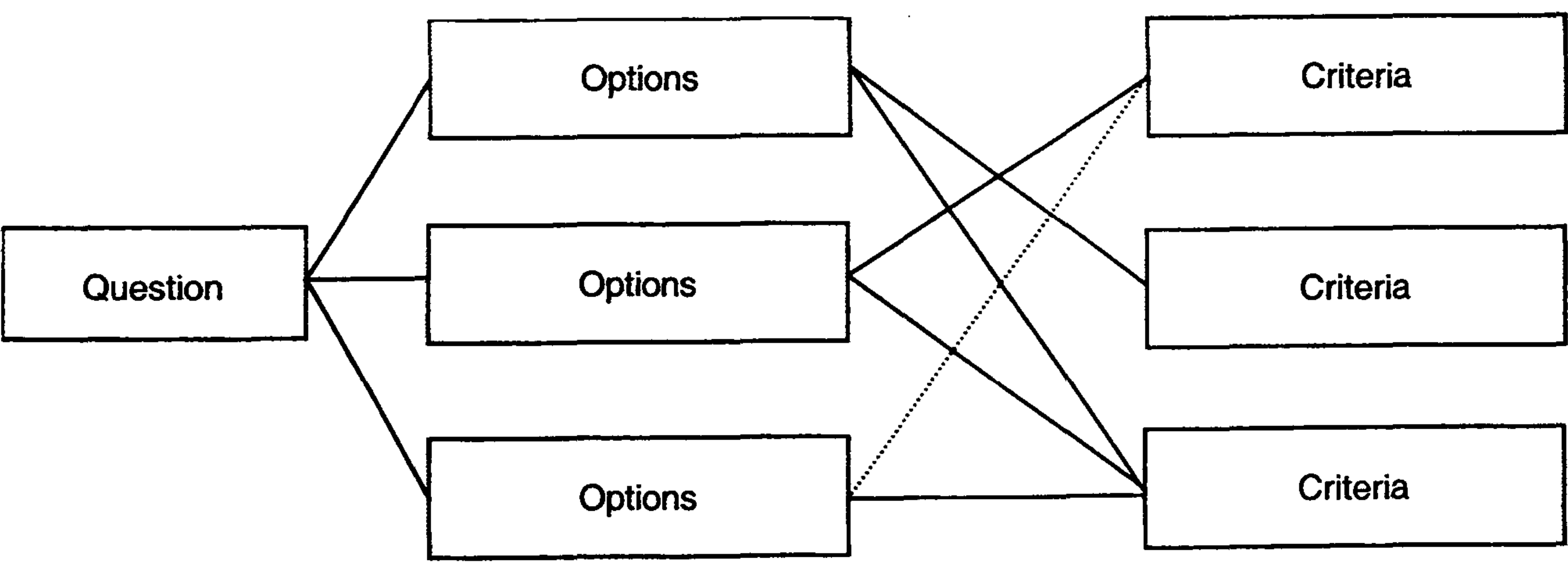


Figure 3.1 QOC Notation

Figure 3.1 illustrates the notation. Options used to satisfy a posed question are presented. The criteria used to select the most suitable option are then considered where positive relationships between options and criteria are represented at solid lines and negative relationships as dashed lines.

### 3.4.2 *Method*

The interaction designer was very familiar with the prototype development software (Macromedia Director v5.0) and had trained as an industrial designer. He had had no training in HCI or human factors but was very experienced in developing interactive product prototypes. In the early stages of the project, a range of interaction style concepts were discussed using pen and paper against the user requirements and wishes of Electrolux but also in terms of their prototyping or coding viability. The designer worked mainly autonomously with design discussions only being called for coding problems or design clarification. Frequent informal progress discussions were held every two to three days. Periodic review meetings were also held with Electrolux to inform them of progress.

Published design guidelines (Appendix 13.1) were readily available. The designer was shown how to use the database using a keyword search method. The designer was told that he could use any of the resources if required or ask if he was unsure where to look.

The prototypes were developed using Macromedia Director, Adobe Freehand and Photoshop software using a Macintosh Power PC 7500/100 with a 22" monitor. Development of all the prototypes was carried out concurrently which helped ease code duplication. Once all four prototypes were complete, keystroke-logging facilities were embedded to support usability evaluation trials (see Chapter 4). After project completion, the designer wrote-up his diary more formally for analysis (Appendix 13.2 - Interaction designer's diary).

### 3.4.3 *Prototype solutions*

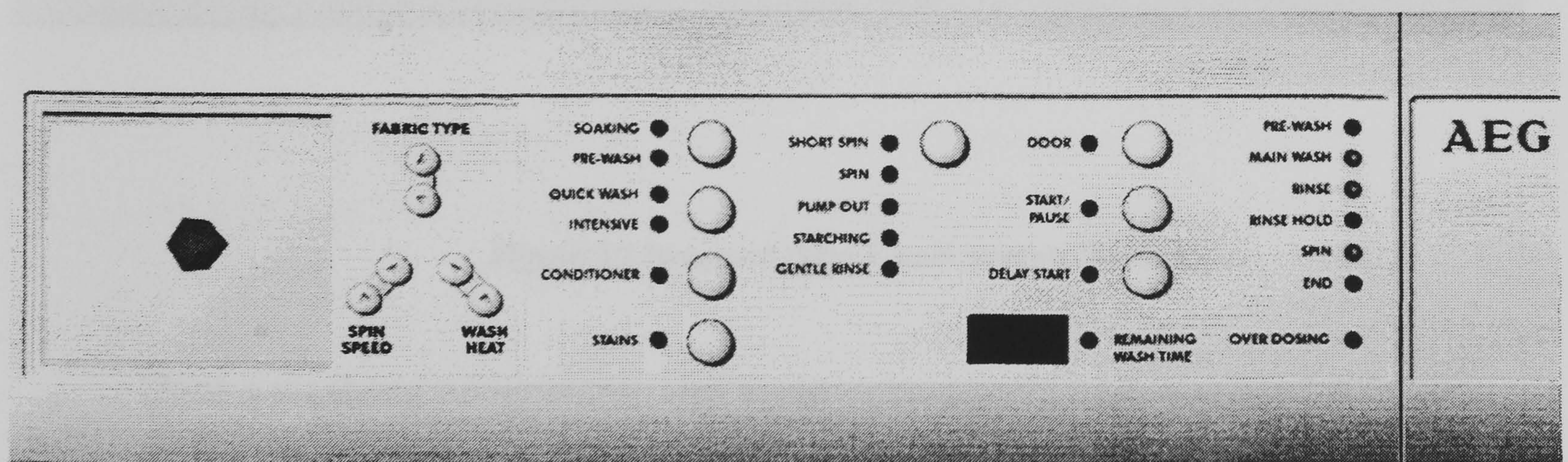
The design and development of the four interactive prototypes took approximately five months from initial concepts to a point where they were sufficiently robust and complete for usability evaluation studies to take place. Descriptions of the final prototype interfaces are presented below so that a



more informed understanding of the design process and novel interfaces developed can be gained. Three novel interfaces were developed based on the AEG Lavamat 6955 plus a prototype version of the existing AEG interface.

### *Animated Object Display (AO)*

This interface used graphic representations of three washing programming parameters (spin speed, fabric type and temperature wash) as an animated object which is altered using conventional toggle switches. This interaction style introduced several novel concepts. Firstly, one graphical object was used to represent three washing programming selections. This was designed to explore if users can use and distinguish different programme selections using a single animated object displaying three programming parameters: colour of the object representing temperature (blue to red), rotation speed of the object representing different spin speeds and object shape describing fabric type.



**Figure 3.2 Animated Object Display**

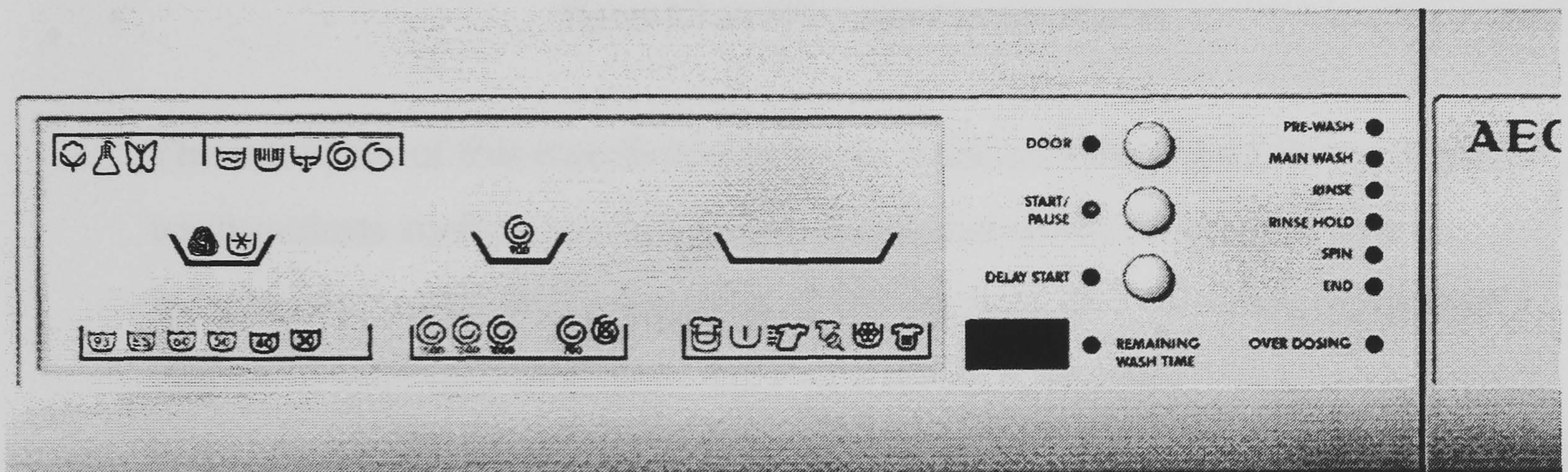
The AO interaction style introduced the concept of 'vague' or imprecise settings where numerical values were not used. The intention was to map functions to more natural, real-world attributes. This allowed a more natural



interaction, for example, using a blue object for a 'cool wash' combined with a slow moving action for a 'slow' spin.

### *Drag and Drop Touch Screen Display (DD)*

This interface exploited the use of spatial movement using a finger controlled 'direct manipulation' based interaction style. Icons represented washing functions, which are dragged and dropped into washing baskets to 'build' a washing programme. This interaction style explored how transferable well-proven Graphical User Interface (GUI) interaction style behaviours can be used on consumer products.



**Figure 3.3 Drag and Drop Touch Screen Display**

Once a washing programme and temperature have been dragged into the first washing basket, a default spin speed 'jumps' into the second washing basket, but can be altered by dragging an alternative spin speed icon if required. Other washing options such as the use of conditioner can be selected in the third basket.

### *Auditory displays (AD)*

This prototype interface exploited the use of audio displays. Visual cues were avoided as much as possible to ensure users were relying on audio cues



during the subsequent user trials. For example, an audio display used a 'whizzing' sound, which altered in pitch, to represent the selection of different spin speeds. A review button was also provided to represent the washing programme using a series of auditory sounds.

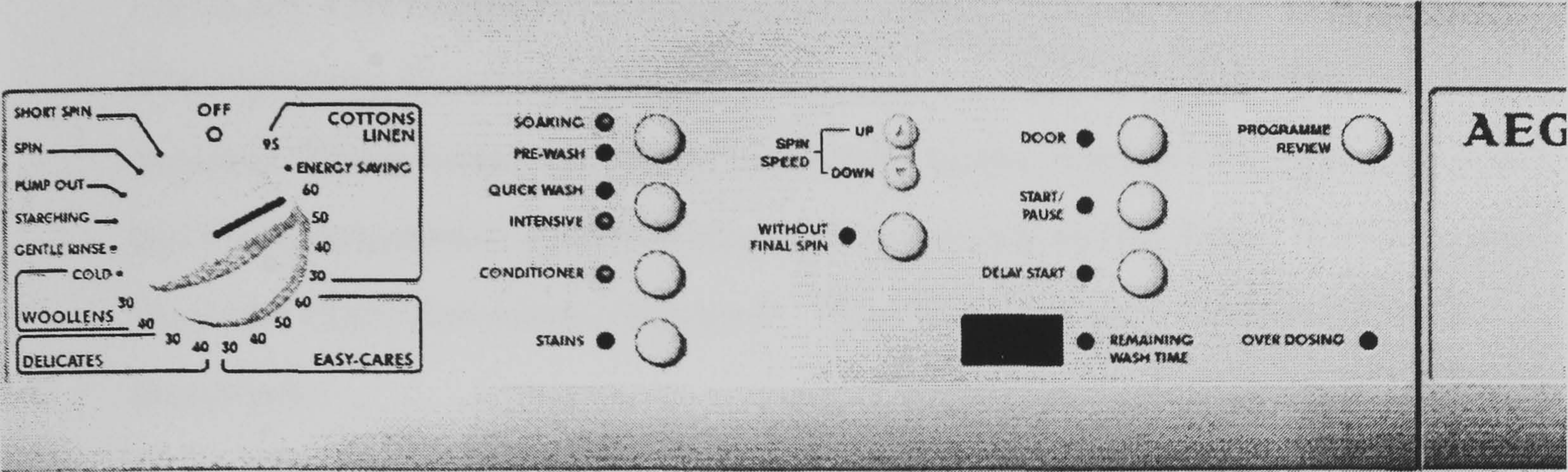


Figure 3.4 Auditory based screen Display

The intention of this interface was to see if the addition of auditory displays for functions such as washing cycle status and control feedback would increase the user's understanding of the product, also to measure the user's degree of tolerance and preference of auditory displays.

*Conventional control washing machine panel (CC)*

A 'contemporary' washing machine control panel was also developed to act as a benchmark and control condition during the evaluation studies in Study 3.

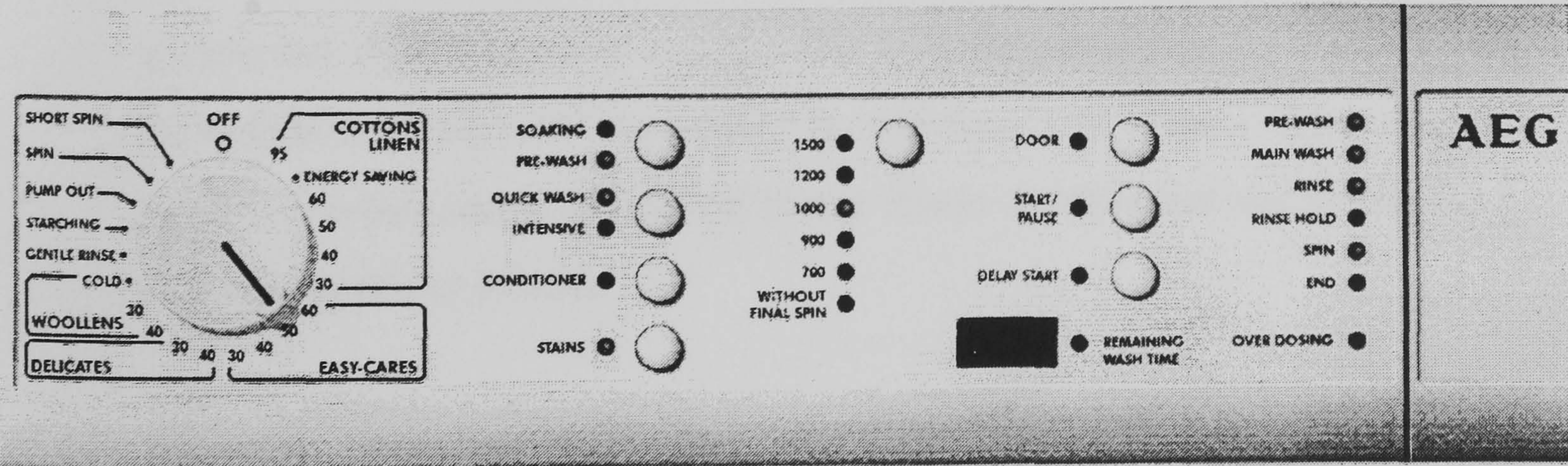


Figure 3.5 Control Screen Display



The functionality of this prototype performed in exactly the same manner as the existing model with a rotary control, illuminating controls, button depressions and associated clicks.

### 3.4.4 Results

The diary was compiled retrospectively rather than on a day-to-day basis. The reporting therefore suffered from brevity with only key events being recalled. There were occasional references to the need for design guidelines but these comments were not discussed in detail in the diary. Direct quotes from the diary have been italicised. The diary reported on three key design problems:

- design and behaviour attributes of icons
- selection of suitable metaphors for two of the prototype variants
- selecting and designing auditory icons.

#### *Behavioural attributes of icons for the DD interface*

This design problem was written about in more detail probably because it offered more programming challenges. Different approaches to problem solving were reported particularly where solving one design problem was 'contingent' on a previous design solution. High level problems were first resolved followed by associated detailed design issues. For example: '*Once an icon was able for selection and was able to be dragged across the screen, I was posed with the problem of determining its behaviour*'; or '*it was only on trying to perform a certain [interaction] task that it became evident that certain [design] decisions needed to be made*'.



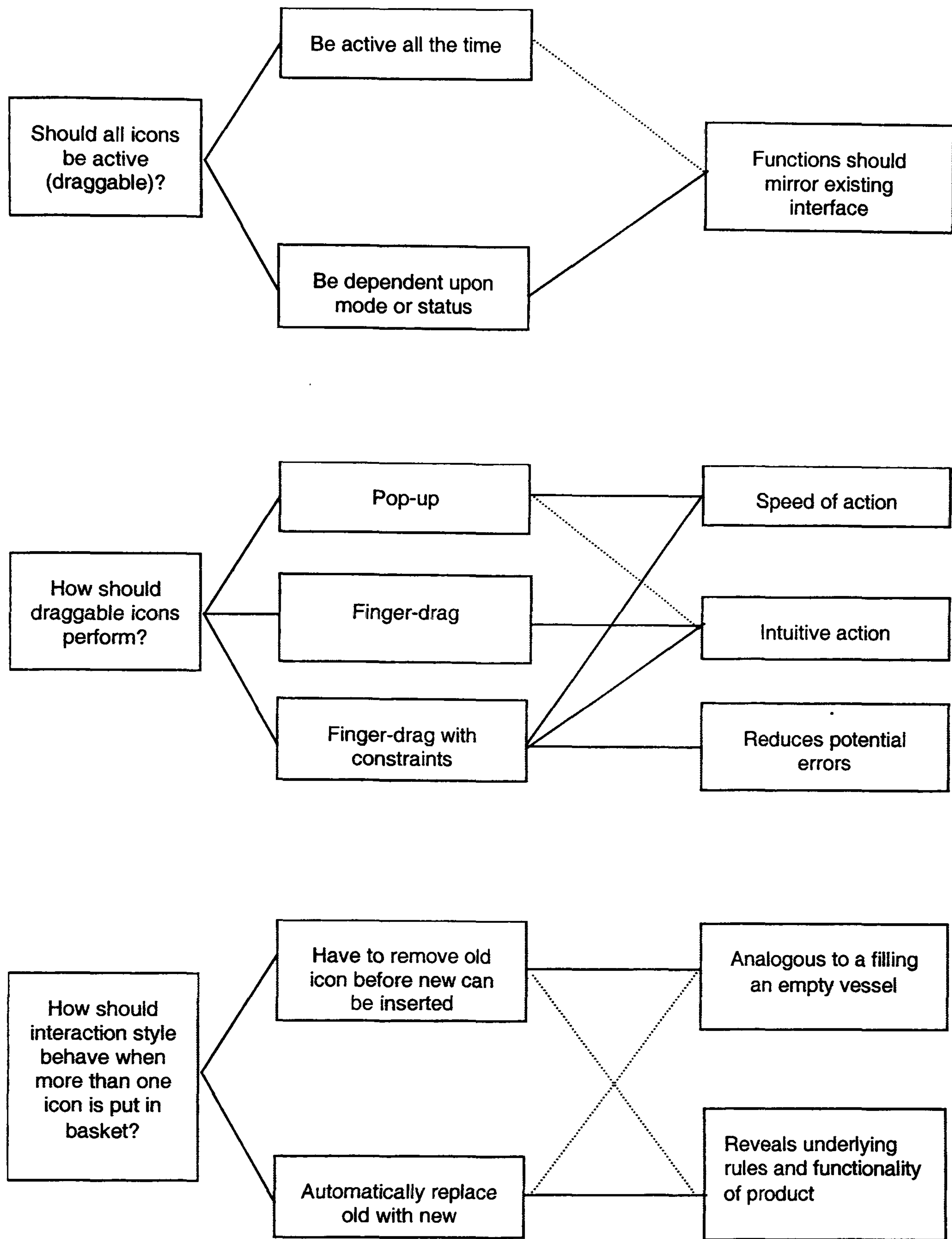


Figure 3.6 QOC analysis of decision making path for icon dragging behaviour

Using QOC, it can be seen that possible solutions to icon dragging behaviour quickly began to diverge. The designer framed the problem as incremental and linear, *'The way I tackled this long list of decisions was to deal with them as each one came up.'* The problem was de-constructed by dividing icon selection and dragging into a series of temporal steps towards the goal of correctly configuring all three washing baskets, see Figure 3.6. This was first done by dealing with the question of when icons should be made 'draggable'

(selection task), followed by how icons should perform while being dragged (action task). This concluded by considering interaction behaviour if more than the maximum number of icons were inserted into a washing basket (error recovery task). The formation and refining of this novel interaction style has been achieved by considering a narrow range of possible and contingent interaction options.

A high-level design question was never formally posed (what types of behavioural attributes could icons have?) but was broken down into three lower level, task-related questions. In examining the first question, only two options were considered, 'be active all the time' or 'be dependent upon mode or status'. The only criterion considered, to resolve these options, was that the selection of icons should be the same as the existing interface. As the existing interface only permitted certain selections depending on mode options, it was decided that the prototype should mirror this, thus resolving the problem by the designer.

For the second question, (how should draggable icons perform?) three options were considered including 'pop-up' (when icon is selected it disappears and re-appears at target) 'finger drag' (icon moves with finger's contact point on display) and 'constrained finger-drag' (icons moves with finger contact point but only towards target). Criteria for selecting these options were based on the designer's anticipated user/task reactions to these different behaviours. Task-based criteria appear to be common grounds for justification. As MacLean and McKerlie (1995 p189) state, 'the task itself acts as part of the argumentation justification, justifying the presence of the criteria derived from it'. Similar evidence can be found in the third problem of how to deal with inserting too many icons into a basket. Here a simple assumed mental model was used to resolve the problem by anticipating a range of potential usage scenarios, *'This would seem logical, as in real life if a container is full, you have to remove some of the contents before putting something in'*. Consideration was not given to supporting other possible options such



as being able to select icons in a different order (for example selecting icons from right to left rather than left to right) or recovery from other types of potential errors (for example making a poor programme selection).

The designer did not report on the use or inclusion of any design guidelines during this particular problem solving activity to help support the generation of design questions, selection of options or consideration of criteria.

Other problems were resolved using prior knowledge of common interaction dialogue such as 'greyed-out' unavailable options, *'The way a typical computer interface does this is to 'grey-out' icons, so I chose this method of communicating information.'* There were also examples of using intuition such as, *'I made an inherent (sic) decision to visually highlight the basket by way of a black border which appeared around it, every time the icon rolled-over the hot area [active part of the display].'*

#### *Selecting metaphors for the DD and AO interfaces*

Two examples in the diary illustrated the use of 'metaphor' to improve the usability of the novel interaction styles. Strictly, the term metaphor referred to in the diary for the first example was not correct. A washing basket metaphor is proposed, *'I tried to think of something that was synonymous with the task in hand. Using graphical 'washing baskets' for the icons to be placed in for washing seemed to be a suitable metaphor.'* The metaphor is tenuous with little mapping between real washing activities and washing machine functions and related more to visual imagery rather than providing useful affordances about how to use the interface. Furthermore, once selected there was little motivation to explore alternative 'metaphors', *'this was the first generation idea which was simply gone along with'.*

The second metaphor introduced attempted to use more 'natural' graphical representations for three programme parameters (temperature, fabric type and spin speed). Colour and rotational speed were chosen for their 'obvious'

metaphorical mapping with temperature and spin speed. Four fabric types were mapped against a morphing object. This object morphed, in four incremental steps between a cube and a sphere. Arguments were put forward for the selection of these 'abstract' objects, for example a cube was chosen *because 'its hard lines, and corners seemed to well represent cotton (in an abstract way) by expressing its hard wearing, tough nature.'* Clearly these shapes do not have any direct relationship with fabric types; the rationale for using them was more likely driven by constraints in 3D modelling than by seeking good representational mapping.

Again, the high level question 'what type of metaphor is most suitable?' was not posed. One metaphor was proposed along with a single option and criterion statement, see Figure 3.7.

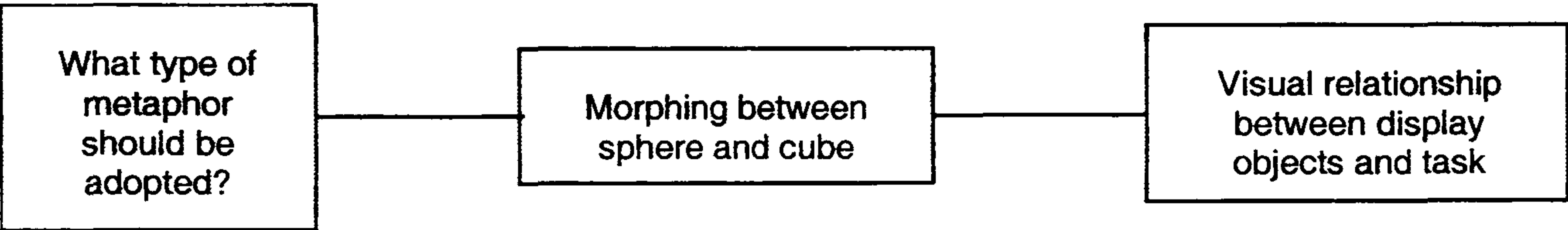


Figure 3.7 Single option and criterion for problem solving

The decision making process has been dictated by simplicity in object modelling rather than the argument explicitly presented in the diary of using a natural relationship between display objects and a washing task. Again, the inclusion or consideration of interface guidelines was not reported.

*Selecting and designing auditory icons*

This proved to be a design problem that was supported by design guidelines - where auditory icons should be used and their duration. However, few guidelines existed on types of sounds that should be used. This resulted in using intuition, *'in selecting sounds to represent certain parameters, [it] involved*



*not much more than guess work.*’ Diary evidence for this design activity was very brief and therefore it was not possible to undertake a QOC analysis.

#### *The use of guidelines*

Despite being explicitly asked before writing the diary to report on the use of guidelines, the diary had no direct references to when and where guidelines were used other than to suggest guidelines could have been useful for icon ‘dragging’ behaviour. Guidelines, however, *were* used during the design process as these were discussed and implemented during design review meetings. These included how icons should be designed (Galitz 1993, Gittens 1986), size of icons (Sears and Shneiderman 1991, Sears *et al*, 1993) and the advantages and disadvantages of different types of icon contact behaviour and corresponding target sizes around icons (Potter 1989).

Design characteristics of auditory icons were also found (Blattner 1989). The lack of design guidance was discussed and the designer briefly mentions where guidelines would have been useful. Table 3.1 provides a summary of the guidelines that were discussed at design meetings and implemented during prototyping and those that would have been useful, had they been available.

Prototype variant	Guidelines found in the database	Guidelines required but unavailable during design process
Drag and drop	Icon design (graphical representation)	Behaviour of icons - when dragging around the display
	Size of icons	Choosing a suitable metaphor
	Contact behaviour	Syntax procedures for dragging icons
	Target size around icon	Default settings
Animated object	Some information on the graphic representation of complex information	Different methods of graphically representing data
		Effectiveness of representing information by colour, speed of movement, shape etc.
		What type of abstract information can be remembered by users
Auditory displays	Sound composition and characteristics	Selection of sound types, appropriateness
	When they should be used	Memory retention for auditory displays
		Discriminability between different auditory displays

Table 3.1 Guidelines found and unavailable during prototyping

A lack of references to guidelines, or the use of them, in the diary meant that the explicit analysis of applicability criteria was not possible. The diary does, however, indicate the designer’s obvious preference for gaining knowledge through mini-experiments, for example *‘By way of experiment, I amended the enclosed boxes by making them open-ended’*.

3.4.5 Discussion

The study centred around one designer. A longitudinal study involving a larger sample of interaction designers working on live design projects in real time would have been almost impossible to implement given the obvious constraints in intervening in commercial design activity while evolving novel



interaction styles. The study attempted, as far as possible, to replicate a real-world design problem. Applicability was deliberately not experimentally controlled which had been the case with studies found in the literature. If this study had controlled the usage or access to design guidelines, the behaviour of the designer would have been influenced by these controlling factors. It was important that that design activity was carried out as naturally as possible. Therefore, a simple, but important, decision was made early not to intervene, as far as possible, in the designer's decision making process or to force the use of interface design guidelines. Key design objectives were set out, but beyond that, the designer had to decide how best to achieve those design objectives. This could only be achieved by conducting an 'off-line' design project guided by Electrolux.

Persuading the designer to keep a diary was difficult; he was not self-disciplined enough to record design activity on a regular basis. The resulting diary was therefore more a summary or recollection of key design activities over a five month period and was relatively brief. Nevertheless, it did provide a personal and subjective viewpoint to the interface design project and permitted some useful analysis. An exit interview was considered but the designer was, by this time, all too aware of the research objectives to make any findings valid. This awareness to a more usability centred approach was evident in the diary. Terms such as 'mental model' and 'metaphor' were used; concepts the designer was unaware of before taking part in this study.

Diary references to the use of guidelines were rare. One of the reasons for this may have been a lack of coercion to consider guidelines. It quickly became apparent during design meetings that the designer was not keen to use guideline resources available to him. To encourage their use, during some informal design meetings, attempts were made to seek out relevant guidelines. Unfortunately, few guidelines mapped directly to the prototype applications under development. This problem concurs with evidence found

in the literature review. This only reinforced the designer's belief there was little to be gained from seeking out guidelines.

A more important reason for lack of adoption appeared to be the designer's propensity to use his own knowledge and experience. The diary is full of references to this type of problem-solving method. This evidence supports findings from other studies such as Klein and Brezovic (1986) and Hammond *et al* (1983) where experiential knowledge was preferred to guidelines. This preference for experientially based knowledge impinges strongly on the applicability criteria, *relevance to the proposed target audience*, suggesting that if the uptake of guidelines was to improve, they must be part of the experiential process and not an addendum. Passive, non-bespoke guidelines were not perceived as having significant cost-benefits to warrant acceptance.

For guidelines to succeed they must quickly prove their worth in providing support at critical and ephemeral decision making moments. These critical points are where the applicability of guidelines could be improved. Existing guidelines did not provide options (different solutions to a problem) or criteria data (qualifying statements for or against options). By adopting a more solution-focussed approach, guidelines could be made more effective.

Also of concern was the quality of tacit knowledge the designer used or gained through experiential activities in the project. The diary provided evidence that simple user/task models were applied to solve interaction design problems, but, more importantly, there was evidence that design problems were sometimes framed around programming problems before addressing usability problems. A good example of this was the morphing object, where simple geometric shapes such a sphere and cube (which were simple to model) was rationalised as being a good representation for abstract concepts (wool and cotton). The designer had mentally 'run-through' the ease of the programming task before offering it as an interaction design



solution. Other interaction styles were not considered as programming resources invested in this metaphor accumulated.

Another example of the use of poor experiential knowledge, was the adoption of metaphors in the design process. The diary suggested the designer was aware that the use of metaphors was good practice but had limited understanding of how these should be implemented. In the case of the 'animated object', the designer had not considered if the overall design concept of three merged metaphors was usable. It was only resolved at the individual metaphor level. A holistic viewpoint was not considered when these metaphors were combined. Again, it was thought that guidelines might be improved by 'packaging' this form of design data within design guideline options. Rather than explain what a metaphor is and how it is of benefit, the applicability of guidelines could be improved by providing a range of metaphor examples and explain how they could be used.

The diary recounted problem solving activity through progressive disclosure. As one solution was implemented, it revealed a further set of design problems, which in turn were resolved. This strategy may be due to several reasons. First, the designer was relatively inexperienced at interaction design (although very experienced at interaction visualisation) and therefore lacked a repertoire of design routines that he knew to be successful. Each problem needed to be resolved at first hand and success could only be measured once complete, resulting in contingent problems not being anticipated. Second, design problem solving was exacerbated by the level of novelty within the proposed interaction styles implemented within each prototype. Finally, due to the novelty of the interaction styles, a conceptual design brief was given to the designer. The conceptual ideas could only be tested once the coding for the novel interaction styles had been proved. Presenting the design brief in this way may have encouraged a more step-by-step problem solving approach. Many of the proposed interaction styles had few design precedents to aid judgement.

### 3.5 Conclusions

Balancing trade-offs between general and specific guidelines in Study 1 was more challenging than had been anticipated. In many respects the CPT guidelines did not overcome the specificity problems already identified in the literature review. However, the key objective of this development study was to gather relevant content for further evaluation studies.

Findings in Study 2 supported previous studies in the literature review. This study found that despite the availability and use of guidelines, they competed poorly with experiential knowledge. The designer was reluctant to add to his design repertoire by using guidelines. Analytical evidence from the diaries found that guidelines, if framed in an appropriate manner, could potentially be made more relevant by supporting option and criteria based decision making. This study reinforced the notion that guidelines must be an integral part of experiential learning activity. As the study was only based on one designer it would be dangerous to generalise the findings too far. However, the findings do support other related HCI studies. More importantly, this study provided an opportunity to witness, at first-hand, a designer making a series of interaction design decisions. This allowed potential avenues to improving guidelines to be considered. Guidelines that offer effective support must suggest alternative design solutions and their consequences rather than just provide device level guidelines.

These findings were used in producing a set of draft CPT guidelines (Appendix 13.4) for evaluation (Chapter 5). However, before these draft guidelines were developed and evaluated, the third strategy for developing CPT guidelines from empirical evidence (which ran in parallel to Study 1 and 2) was carried out. This is described in the following chapter.



## 4 Producing CPT design guidelines from usability evaluations

---

### 4.1 Introduction

In this chapter Studies 3 and 4 are presented, forming the third guideline development strategy (see Section 2.8). The primary purpose of these studies was to test the feasibility of *proactively* developing new design guidelines specific to CPT. In Study 3, three 'high level' novel prototypes and one control prototype washing machine interfaces were evaluated to produce new CPT guidelines. Problems in prototype design prevented the evaluation studies from being as sensitive and productive in producing guidelines as was hoped. For this reason, Study 4 was introduced in an attempt to develop guidelines using a different approach - 'low level' prototyping.

### 4.2 Aims and objectives

- To test the feasibility of proactively developing guidelines from novel prototype evaluation studies

### 4.3 Study 3 - Evaluation of high level CPT prototypes

Prototypes can be developed to different degrees of fidelity and can be 'high level', where all or most of the functionality and often the form of the interface is fully represented, or 'low level' where only a local or specific element of the interface design is modelled. Many of the advantages and disadvantages of high and low level prototyping have been reported, for example Rudd *et al* (1996) offered advice on the applicability and resource implications for both approaches. The purpose of this study was to investigate how feasible it would be to develop guidelines from a usability study using 'high level' interface prototypes designed with this intent.

The three high level novel CPT prototypes produced in Study 2 were used for the usability evaluation study. The novel prototype interfaces have been

previously described in detail in Section 3.3.4. The 'animated object' (AO condition) prototype used graphic representations of three washing programming parameters (spin speed, fabric type and temperature wash) as an animated object. This type of interaction style had previously been used to represent aircraft flight information (Wickens and Andre 1988). The aim was to develop guidelines to support the use of 'vague' settings and presenting multi-function representations within a single object. The drag and drop (DD condition) was designed to produce guidelines related to the usability of a finger controlled 'direct manipulation' interaction style. While the intention of the third prototype, an auditory display prototype (AD condition) was to develop design guidelines related to using auditory displays to express product functionality.

#### **4.3.1 Method**

An evaluation study was designed to ensure that as much design data as possible could be captured from the results. To do this, a list of design questions was first formulated (in effect the obverse of design guidelines) many of which corresponded to unanswered design questions posed during prototype in Study 2. To assist further in devising questions with adequate scope, questions were posed at different abstracted interaction levels, defined as task, semantic, syntactical and device levels based on Moran's interaction model (Moran, 1981). A table for each condition was drawn up with questions posed at each interaction level (Appendix 13.3 - Study 3 experimental material). Table 4.1 provides an example question illustrating the type of data gathered to obtain design data, followed by the intended analysis method.



Questions posed (at each level of interaction)	Data capture	Data analysis
How usable do users think auditory displays are?	Questionnaire using subjective rating scales, comparative analysis between auditory displays	Wilcoxon test or t test: compare rating with neutral answer

Table 4.1 Example question with corresponding capture and analytical methods

Four experimental trial tasks were designed to assist in answering the experimental questions, these were:

- Procedural tasks - where subjects were given precise instructions, for example, *set temp to 60 degrees, select wool setting....* Variants of these instructions were presented at set intervals during the user trial to measure, for example, changes in learning times.
- Narrative tasks - where subjects were presented with a washing problem without any specific instructions on how to complete the task, for example, *you need to get a woollen blanket clean in two hours....*
- Amendable tasks - where subjects were deliberately asked to change their goal during task activity to test task competency.
- Invalid tasks- where some tasks could not be successfully achieved. These were set to test completeness of task competency.

Variant tasks were repeated to measure changes in performance over the trial period. This created eleven tasks overall. Dependent variables were then considered. At the task level the dependent variables were performance (speed and accuracy), perceived usability, learnability, preference and quality of mental models. At the semantic level the dependent variables were comprehension of interface objects (icons, sound, animated object displays) and perceived efficacy (quality of feedback). The syntax level dependent

variables were accuracy and comprehension of the steps to carry out the tasks. Device level questions were not considered due to the lack of tactile fidelity between the software based prototype interface and the existing hardware product interface.

4.3.2 Experimental design

Due to possible learning effects and trial duration for each condition, a 'between subjects' design was chosen using matched participants between the three condition groups. Ten participants were used in each group (see Table 4.2). All conditions contained 11 different tasks (4 procedural, 4 narrative, 2 amendable, and 1 invalid). The same tasks were repeated for all conditions except for one amendable task in the AO condition, which was impossible to do due to the syntax of the interaction style. This was replaced by an additional invalid task.

Participants	Conditions
Group A	Drag and Drop and Control
Group B	Auditory Display and Control
Group C	Animated Object and Control

Table 4.2 Experimental conditions

Order effects were dealt with by counterbalancing conditions and task order. However, counterbalancing task order had to be limited as some tasks were conducted at fixed points during the trial. For example, tasks measuring learning time needed to be evenly distributed throughout the trial to be meaningful.

Evaluation methods

Subjective measures for control and experimental conditions were gathered through two questionnaires (an example is provided in Appendix 13.3 - Study 3 experimental material). In each questionnaire, 21 questions were



posed using a 9 point likert scale, measuring learnability, semantics or perceived meaning and understanding, syntax or knowing what to do next, layout of the interface overall reaction. Using experimentally proven questions by Hix and Hartson (1993), Pare and Elam (1996) and Hill, Smith and Mann (1987) ensured validity of the questions. A second questionnaire was given after completing both user trials to establish comparative benefits and problems. This was a 12 question, 9 point likert scale questionnaire used by Hubona and Blanton, (1996) measuring the participant’s level of confidence for both conditions and also the best and worst aspects of the interfaces, based on questions used by Hix and Hartson (1993). Memorability and understandability were measured through recall and recognition tasks on the different types of display objects, icons in the drag and drop condition, auditory displays and different display states of the animated object (an example is provided in Appendix 13.3 - Study 3 experimental material).

Condition	Recall and recognition tests
Drag and drop (DD)	Free response to 10 icons (recall task)
	Limited response (10) to 13 icons (recognition task)
Animated Object (AO)	Recognition to set animated displays
	Subjective rating on the viability of a animated display
Auditory display (AD)	Recognition to 10 auditory displays (13 washing programme elements listed)
	Ordering 3 sounds in quantitative terms (using pitch)
	Recalling and describing washing programme after hearing a series of auditory sounds sequentially

Table 4.3 Recall and recognition tests for display objects

Table 4.3 summarises the different types of recall and recognition tasks presented. In the DD condition the first test required the participant to *recall* any response to a set of icons and in the second test the participant was

required to *recognise* washing programme elements against 13 icons presented. Extra icons were presented to reduce the danger of participants selecting the icons using an elimination strategy. A similar design was used for the AO condition with the exception that in the second stage of the AO test, the participant had to rate how confident they felt that a particular animated object would work or not.

Again in the AD condition participants were provided with a list of 13 different washing programme elements and then listened to 10 different auditory displays and asked to match them. In the second part of the test a series of three 'family' sounds were presented indicating different quantitative measures, for example high medium and low spin speeds presented in a non sequential order (e.g. high, low, medium). These had to be written down in the order in which they were presented. The final part of the test required the participant to listen to a 'composed' washing programme using auditory displays only and then write down what they thought it was.

To understand the types of mental models participants were forming, a fourth data gathering method known as 'teachback' (Van der Veer, 1990) was introduced. This method identifies and classifies mental models developed by the user using four abstracted levels of interaction, task, semantic, syntax and device, based Moran's Command Language Grammar (1981).

Participants were asked, after completing the experimental condition, to describe to a fictional 'target' reader one of the experimental tasks they had just completed, using their own preferred mode of representation.

Participants were asked to leave instructions on how to set the washing machine for another specified user who has never used the interface before. They were told that they could describe the task in any form they like either as written or diagrammatic form. At this stage, they were not permitted to refer to the interface or any task or help cards used in previous parts of the study. The teach back protocols for each participant were then coded in



terms of completeness and accuracy. Three evaluators were employed to code independently all the protocols, which were subsequently ratified for inconsistencies by checking each other's coding scripts.

A data logging file was set up to measure planning and preparation duration, task duration and keystroke activity. An activity report was produced using the logging file. This required some manual analysis to obtain performance measures.

### *Participants*

Thirty participants were recruited from the Teesside area via advertisements in local papers and posters in local libraries. Following telephone interviews, participants were matched against age, gender, type of automatic washing machine owned, experience in using a computer and educational level to reduce condition variability. Matched participants were divided into three groups of ten, each group evaluating the DD, AO and AD interfaces against the control condition CC. Participants were paid £10 for taking part in the study.

### *Materials*

The prototypes were developed using the Macromedia Director v5.0 software using a Macintosh Power PC 7500/100 with a 21 inch colour monitor.

### *Procedure*

User trials were carried out by first explaining the purpose of the trial and to allay concerns. Each participant was briefly shown how to use the washing machine interface. After this, participants carried out the 11 washing programme tasks in condition 1 (control or novel interface). Participants were given a task card containing instructions of a washing programme to be entered. Participants were allowed to refer to task cards and legends at any point and were also instructed that if they were unable to complete a task they were permitted to ask the experimenter any questions relating to the

interface. Apart from this, the experimenter was not permitted to help the participants. Once completed a questionnaire was presented measuring subjective views and object recall and recognition skills, followed by a 'teachback' activity.

Condition 2 then followed, using a similar structure to the first condition; participants carried out 11 tasks followed by a questionnaire and teachback activity. The trial concluded with a comparative questionnaire on the differences between the two interfaces (examples are provided in Appendix 13.3 - Study 3 experimental material).

### 4.3.3 Results

The results are presented in terms of the evaluation methods used across all condition groups.

#### *Post interface condition questionnaire*

The post condition questions, which were known to measure similar subjective dimensions, were tested for correlation between them. An analysis of the correlation within the themed questions showed that two questions from the 'learnability' questions (Spearman  $\rho = .577$ ), and two from the 'semantics of the interface' questions (Spearman  $\rho = -.831$ ) correlated significantly. The sum of the scores for each of these pairs was calculated and subsequently used. All five questions concerned with 'overall reaction' (Spearman  $\rho$ s between .444 and .812), two questions concerned with 'fun' correlated significantly (Spearman  $\rho = .869$ ). The sum of the scores for each of the significantly correlated subsets of questions was calculated and subsequently used.

Rating whether the novel interfaces were different from the control interface on these measures was then conducted using a Mann-Whitney U test comparing subjective measures between the first interface condition presented and the second interface condition. Significant differences were



found between the control and experimental conditions with Question 1 (*I know what all the different controls can do*), Question 10 (*It was very quick to learn*) and the combined Questions of 2 (*I know what washing programmes are available*) and 10. When comparing the interfaces on the remaining measures, Wilcoxon tests were used to test the control condition interface as rated by participants in either the first or second interface condition against the DD and AO interfaces. When comparing the interfaces on the remaining measures, Mann-Whitney U tests were used testing the control condition interface as rated by all participants against the DD and AO interface.

Participants found the AO interface more difficult to learn (Median = 3) than the CC interface (Median = 7) (Wilcoxon test:  $T = 6$ ,  $p < .05$ ). However, participants thought that after a mistake, the AO condition was clearer (Median = 4), in suggesting what to do next compared to the control condition (Median = 0) (Mann-Whitney U test:  $U = 53$ ,  $p < .05$ ). The DD control panel (Median = 2) was rated as being more complicated than the CC (Median = -3) (Mann-Witney U test:  $U = 36.5$ ,  $p < .05$ ). The control condition interface, to a larger extent, contained all the information needed (Median = -2) compared to the DD (Median = 1.5) (Mann-Whitney U test:  $U = 46.5$ ,  $p < .05$ ) and AO conditions (Median = 4) (Mann-Whitney U test:  $U = 18$ ,  $p < .05$ ).

#### *Comparative questionnaire between conditions*

Table 4.4 presents a summary of clustered responses made by the participants to the question '*What were the best features of the (standard or novel) washing machine panel?*' The number of possible responses for the control condition was 30 as all participants were exposed to this condition, whereas the maximum for each novel condition was 10. These are expressed as percentage figures.



Condition	Best Aspects	Number of responses (%)
Standard control panel (CC)	Easy and fun to use	27
	Rotary dial control	13
	Information/status presentation	13
	Panel layout	10
Drag and Drop (DD)	Easy and fun to use	50
Auditory Display (AD)	The use of sounds	50
	Easy/Fun to use	30
Animated Object Display (AO)	Use of animation and graphics	30

Table 4.4 Best aspects of interface conditions

Table 4.5 presents subjective statements for the worst aspects of the two interfaces participants were exposed to.

Condition	Worst Aspects	Number of responses (%)
Standard control panel (CC)	Too many buttons or functions	13
	Not able to reverse direction of rotary dial	10
Drag and Drop (DD)	Identification of icons	70
Auditory Display (AD)	Distinction between sounds	50
	Too many/irritating sounds	30
Animated Object Display (AO)	A few comments made but without general consensus	N/A

Table 4.5 Worst aspects of interface conditions



In response to the question ‘*What common mistakes would you say you made on the (standard/novel) washing machine panel?*’, participants reported that only with the standard interface were no mistakes made (Table 4.6).

Condition	Common mistakes	Number of responses (%)
Standard control panel (CC)	Reported 'none'	20
	Turning the dial too far	27
	Selecting the wrong temperature setting	12
Drag and Drop (DD)	A wide dispersion of low frequency mistakes	N/A
Auditory Display (AD)	Selecting the wrong spin speed	40
Animated Object Display (AO)	Selecting the wrong temperature	50

Table 4.6 Reported common mistakes for interface conditions

Turning the dial too far can be attributed to the programming of the prototype where in some circumstances there could be a delay between selecting the rotary control and the prototype responding. This caused some participants to ‘hit’ the control repeatedly to gain a response, only to find the control slowly moving past their target selection.

Table 4.7 presents responses to two questions: ‘*what were your initial impressions of the (standard/novel) washing machine panel?*’, and ‘*what changes would you make to the (standard/novel) washing machine panel?*’. Only the most common frequency responses are given.



Question	Comments	Number of responses (%)
Standard control panel (CC)	<i>Initial impressions?</i>	
	It was familiar	21
	'Likeable'	15
	Complicated	9
	<i>What changes should be made?</i>	
	None	33
Drag and Drop (DD)	<i>What changes should be made?</i>	
	Design of icons	70
Auditory Display (AD)	<i>What changes should be made?</i>	
	Spin speed should be visually displayed	60
Animated Object Display (AO)	<i>What changes should be made?</i>	
	Addition of numerical displays	50

Table 4.7 Reported impressions and suggested changes for interface conditions

*Teachback tasks*

An 'overall rating' score was given for each participant. This was a frequency count of the number of interaction levels used within each teach back description. It would have been possible for a subject to respond on all levels for all teach back tasks (4 levels x 3 tasks) and receive a score of 12. Wilcoxon tests were conducted on these data. No significant differences were found in how instructions were explained between conditions.

There was a *tendency* towards fewer levels of interaction descriptors being used while describing the drag and drop condition than the control condition (Mean = 4.5 SD = 1,  $z = -1.859$ ,  $p = 0.063$ ). Also, a greater number of syntax descriptors were used to describe the animated object condition than in the control condition (Mean = 3, SD = 0.75,  $z = -2.023$ ,  $p = 0.0431$ ).



Novel condition	Interaction level descriptors	Correlation between novel and control conditions	p<0.01	p<0.05
Drag and Drop  n = 10	Task	-		
	Semantic	0.86	*	*
	Syntax	-0.37		
	Device	0.22		
Animated Object  n = 10	Task	1.00	*	*
	Semantic	0.59		
	Syntax	0.25		
	Device	0.70		*
Auditory  n = 10	Task	-		
	Semantic	0.30		
	Syntax	0.84	*	*
	Device	0		

Table 4.8 Correlation in interaction level rating scores between novel and control conditions

A Spearman rho correlation was undertaken of subject frequency scores between novel interfaces and the control interface to establish if subjects use the same *level* of descriptors regardless of the condition (Table 4.8). There was evidence of correlation between representation levels at the semantic level for the drag and drop interface and the control interface with a common use of descriptive language. At the device level in the animated object interface and control interface, there was also a more common use of language relating to physical objects. At the syntax level in the auditory interface and the control interface, there was a more general use of procedural language. These results suggest that experimental conditions may



have affected how the control conditions were described as subjects had been exposed to both conditions before undertaking the teach back exercise.

In coding teachback protocols, critical observations of the coding process were made. Protocols were described in a way that someone else would understand (in the form of a step-by-step guide for example) rather than elaborating on how they understood the interface to work. This made coding judgements more difficult. Protocols often did not refer to interface elements, for example in the AD condition few participants described sound types and their meaning. Furthermore, as participants went through protocols, some participants began to omit repetitious parts of the task thus reducing protocol consistency. Some tasks were also too short to make any real analysis for example pressing the 'start button'.

#### *Recognition and recall of interface objects*

For the DD condition, only the icons for '30 degrees' and '1000 rpm' were accurately recalled. Recognition of icons was more successful, however 'prewash', and 'gentle rinse' provided most confusion. For both recall and recognition, icons were identified significantly less than an experimentally imposed target accuracy of 90% ( $p < .05$ ). For the AO interface condition, the accuracy with which participants recognised the settings of all three washing program parameters was fabric type (75% accuracy), temperature (50% accuracy) and spin speed (43% accuracy) versus incorrectly identified settings. Although there was a high level of inaccurate recognition, many of the inaccurate responses were adjacent to correct responses. There were two instances where the extremes of the animated object (in this case a sphere for woollens and a cube for cotton) were confused. The level of inaccurate responses to the AD condition was very high with an extremely wide dispersion of responses. The only exception to this was the woollen sound ('baa' sound of a sheep) which obtained a very high accuracy.



*Data logging*

Total completion times were recorded. A two way ANOVA was conducted on completion times with type of interface (Control and Novel) and type of novel interface (AD, DD, AO) as factors. Significant differences were found between the novel interfaces ( $F = 4.36, p < 0.05$ ). However, there was also an interaction effect between the control and novel interface conditions ( $F = 3.96, p < 0.05$ ). The first condition affects completion times for the second condition. Total times, given in minutes are presented in Table 4.9. All the novel conditions took longer to complete than the corresponding control condition.

	AD	DD	AO
Control interface	32.03	38.55	40.25
Experimental interface	36.54	46.95	62.72

**Table 4.9 Mean task completion times in minutes**

**4.3.4 Guidelines derived from study**

To provide answers to the questions originally set, the results from the usability evaluation were used. These, in turn, allowed design guidelines to be produced. Tables 4.10, 4.11, and 4.12 list the guidelines produced from this study. In the first column the original questions are presented. In the second column design guidelines are presented by answering each question. The third column outlines where the evidence was sought and any additional supporting comments.



Experimental question	Design guidelines	Derivation
How usable do users think auditory displays are?	Heavy reliance on auditory displays needs to be considered carefully. Can be perceived as improving the 'friendliness' of the interface but improving usability, in the short term, is unlikely	Questionnaires  Recognition and recall tests
How well are different washing programmes remembered compared to the control condition?	Avoid the use of abstract auditory displays as recall and recognition of these is very poor. As far as possible use auditory displays that have clear and obvious concrete associations. These are popular and easily recognised but do not necessarily improve usability. If many abstract displays are to be used, these should be designed and evaluated carefully to reduce confusion between them	Recognition and recall tests.  Participants made very favourable comments about some of the auditory sounds selected and enjoyed using the interface because of these despite usability not improving.
Can programmes be learnt quickly and what is the learning behaviour?	Learnability of abstract auditory displays is weak. Differentiation of similar sounds needs to be high to improve learnability or recognition.	Questionnaires  Data Logging  Performance measures were weaker against the control condition
What programming mental model do users have using auditory displays?	Heavy reliance on auditory displays does not improve the construction mental models - certainly in the initial stages of learning	Teach back  Recognition and recall tests
What types of errors are made?	Avoid the use of auditory displays to perform selection tasks	Recognition and recall tests  Participants found it very difficult to accurately select spin speeds
Do users understand the meaning of the composition and differences between auditory displays?	If presenting auditory displays serially, ensure there is at least a 1-2 second gap between displays	Recognition and recall tests
Is there any confusion between auditory displays?	Abstract displays and sounds with similar pitch or content create confusion	Recognition and recall tests
Is the correct level of feedback given with auditory displays?	If possible use visual displays to complement auditory displays	Recognition and recall tests  Data logging
Are all or any of the auditory displays irritating?	Selection and frequency of display needs to be considered carefully, as they quickly become irritating.	Questionnaires
Do the auditory displays behave in a manner expected by the user?	Results from study did not produce evidence to answer this question	Comments from questionnaire did not reveal this as an issue

Table 4.10 Guidelines derived from audio display interface condition



Experimental question	Design guidelines	Derivation
What are the usability issues related to directly manipulating the 'design' of a washing programme using a touch screen interface?	This interaction style is easy to use but ensure that the location, design and movement of objects or icons is intuitive	Questionnaires
How well are different washing programmes remembered compared to the control condition?	Results from study suggested that this interaction style is no worse or better than conventional controls and displays	No significant differences found between control and experimental conditions
Can programmes be learnt quickly and what is the learning behaviour?	This interaction style is regarded as easy and fun to use	Task completion times were slower in the experimental condition, this was mainly due to poor icon recognition rather than the design of the interaction style
What mental model of programming do users have?	This interaction style encourages a 'spatial' understanding of washing programmes. This may help in selecting different programmes	Teach back
What types of errors are made?	Design this interaction style to ensure that good feedback is provided for common mistakes such as: incorrect selection due to inaccurate target selection; incorrect icon selection; or incorrect target selection.	Questionnaires
Do users understand the meaning of the icons?	Ensure that icon recognition evaluation studies are carried out at an early stage of development.	Questionnaires Recall and recognition tasks
Is the correct level of feedback given?	Good feedback is provided in terms of possible actions that can be taken.	Questionnaires
Can the user understand what is possible and not possible at each stage?	Behaviour rules for this interactions style are quickly understood but not immediately intuitive	Results from teachback revealed that behaviour could be accurately described but task completion times suggest that users need time to understand the behaviours of the interaction style
Do the icons move and behave in a manner expected by the user?	The 'dragging' and 'dropping' of icons is not immediately obvious. More visual cues are required to improve this.	See above

Table 4.11 Guidelines derived from drag and drop interface condition



Experimental question	Design guidelines	Derivation
What are the usability factors for setting 'vague' washing programmes?	This interaction style is usable as long as there are large discriminating differences between each setting. Suggest there are no more than 3 settings for each parameter.	Questionnaires Recall and recognition tasks
Can users remember washing programme settings using visual cues?	The display parameters used in this study are not easily remembered. This may be due to the abstract nature of some of the representations used and the lack of discrimination between settings	Questionnaires Recall and recognition tasks
Can programmes be learnt quickly and what is the learning behaviour?	This interaction style encourages a 'trial and error' learning behaviour where error recovery is often implemented	Teachback
What programming mental model do users have?	Difficult to discern from study results	Results from teachback were not conclusive
What types of errors are made?	Ensure that differentiation in object displays is high to reduce incorrect selections.	Recall and recognition
Do users understand the meaning of the different graphic representations for programme selections?	Ensure that object recognition evaluation studies are carried out before implementation.	Questionnaires Recall and recognition tasks
Is the correct level of feedback given?	The interaction style, used in its current configuration, does not provide an appropriate level of feedback	Questionnaires Recall and recognition tasks Task completion times
Can the user understand what is possible and not possible at each stage?	Subtle variation in display representations are not recognised	Results from teachback suggested that dangerous programme selections went unnoticed
Do the controls behave in a manner expected by the user?	Difficult to discern from study results	Results from questionnaires were not conclusive

Table 4.12 Guidelines derived from animated object interface condition



### 4.3.5 Discussion

The results indicated that, overall, no interface condition proved to be comparatively more usable to any other interface condition. Results from the questionnaires given after each condition and the task completion times, however, suggest that the control interface was preferred and quicker to use. This can probably be attributed the familiarity of the control interface.

Many of the usability problems encountered in the experimental conditions can be attributed to using a summative evaluation strategy and deliberately avoiding redundancy in the interface design. Little formative evaluation and few design precedents during the development of these interaction styles, forced design problems to be resolved only at the usability evaluation trials. The level of novelty within each interaction styles prompted more usability design problems to occur during development than was anticipated. As observed in Study 2, design problems were resolved by using assumptions about user reactions. User consultation, at this stage, was not possible, as the functionality of each interaction style needed to be interactive before any objective assessment could be made. Furthermore, problems identified further down the development process were often more difficult to resolve, purely because significant re-coding would have been required.

This resulted in the final prototypes carrying 'gross' usability problems. For example, in the DD condition, icon design was very poor in terms of comprehension, the AO condition presented objects that were too abstract and in the AD condition spin speed sounds were offered without other forms of visual display. This weakened the 'sensitivity' of the evaluation methods to reveal interaction behaviour that would have helped in developing more insightful guidelines. However, if redundancy had been offered by adding conventional interaction styles, it would have been difficult to isolate usability factors directly attributable to the novel interaction styles. In retrospect some of the usability problems seemed obvious, but at the time it



was difficult to gauge what design elements or usability factors would affect user behaviour and attitudes towards the novel interfaces. It was hoped that the evaluation study would reveal subtle usability problems as well as the obvious ones. However, in effect, the gross usability problems clouded observations of other potentially subtle usability problems.

There was also another factor that affected the push towards the completion of interactive prototypes. The original intention was to explore guideline development through a series of low level prototypes, but during the initial discussions with Electrolux there was considerable misunderstanding about the purpose and development aims of some of the low-level prototypes proposed. The marketing department in Electrolux, who were funding this study, was anxious to gain credible, high quality prototypes. Gaining guidelines to benefit the development of future CPT was not as important to them as obtaining 'attention grabbing' prototypes for marketing purposes. This forced the development of high-level prototype sooner than was intended.

The guidelines produced were few and general in tone. Using a traditional formal experimental approach was found to be heavy-handed and did not provide the level of design and evaluation flexibility to produce useful design data. The number of guidelines generated was few and not particularly revolutionary. Attempts to understand the usability of novel interaction styles by isolating them and removing redundancy proved to be counterproductive. Guidelines can only be produced when more is known about the context or environment in which novel interaction styles will be used. The key failure of this study was not integrating or formalising small user tests to resolve occurring usability problems.

Electrolux accepted lessons learnt from this study and a different approach was agreed upon. Resolving usability problems (and thus producing guidelines from them) had to occur when a problem was identified. A series



of formative rather than summative evaluation studies were considered in the hope they would create more effective guidelines by closely integrating design and evaluation activity. A further study was therefore proposed (Study 4) using more informal, 'quick and dirty' formative evaluation methods to develop effective design guidelines based on evolving usability design problems.

4.4 Study 4 - Evaluation of low level CPT prototypes

This study was designed to offer an alternative or counterbalanced approach to Study 3. Table 4.13 summarises the differences in guideline development strategy. The objective of this study was to compare the effectiveness of this approach with the previous study for gathering CPT design guidelines.

Approach	Study 3	Study 4
Evaluation design	Formal design, controlled experiments, testing through statistical analysis, medium sized user groups	Informal design, 'quick and dirty' experiments, testing through subjective comments and frequency counts, small user groups used, reflects a more iterative interface design approach
Evaluation stages	Summative evaluation	Formative evaluations triggered by usability design problems
Interaction styles evaluated	Three interaction styles evaluated concurrently using one iteration cycle	Focus on one interaction style, developed with local iterations
Approach to guideline development	Query based approach based on 4 layer interaction model to achieve device based guidelines	Observation based approach to derive device and 'design process' based guidelines

Table 4.13 Comparison of experimental design between Studies 3 and 4

Design and evaluation activity was integrated more closely using formative evaluation methods. Quick and dirty evaluation methods were introduced taking no longer than a day to initiate and complete. In order to do this, low level prototypes were developed describing a single functional component of a novel interface. Evaluations involved small participant groups (5-6



subjects). Evaluation studies were not designed or intended to be experimentally robust. They were intended purely to provide a designer with some insight into the potential usability problems that might exist with a particular novel interaction style. Speed and quick feedback were the critical factors.

### **4.4.1 Methodology and results**

In conjunction with Electrolux, it was decided to develop the 'drag and drop' interface, which appeared to be the most likely candidate for commercial development. The UK design group was working on microwave interfaces providing an opportunity to explore this interaction style further using a different product platform. The previous study had revealed that participants enjoyed using the drag and drop interface but were hampered by poor icon design.

Interface design began in a similar way to Study 3. Ideas and concepts were discussed with the interaction designer and Electrolux using sketched ideas. The interaction designer used in Study 2 then began on the prototype development work with instructions to produce low level options or variants when a novel interaction style design problem occurred. Options were then generated so variants could be evaluated using 'quick and dirty' methods. During the interface design work, two small incremental development steps (Icon selection and Icon behaviour) were introduced thus reflecting a more 'local level' responsive usability design and evaluation process.

#### *Icon selection*

The first stage consisted of developing and refining icons for the proposed microwave interface. A questionnaire with sketch sheets was distributed to 48 non-paid students of Teesside University (29 design and 19 non design) who were asked to draw their own depiction or representations of 12 listed functions currently available on an Electrolux microwave.



Students were also asked to illustrate possible concepts for animated icons by drawing 'storyboard' frames. These drawings and images were then classified according to their referent characteristics to the relevant microwave function. Using simple frequency counts, three to six sample icons were then selected for each microwave function reflecting the most stereotypical or common icon depiction from students. Selected icons were then presented to a different student group ( $n = 41$ ), who then subjectively rated and annotated each icon for meaningfulness. Again simple frequency counts and logging of comments were used to select 12 final icons. Eight of these were selected to reduce prototype development time; examples of these icons are presented in Figure 4.1.



**Figure 4.1 Icons selected through participant consultation**

### *Icon behaviour*

In the second stage of this study, the selection, dragging and behaviour of icons needed to be resolved. Proposed solutions were developed by the interaction designer offering four different interaction style component variants to accomplish the same task. Variants were evaluated by paired comparisons with all the possible permutation pairs being presented. Six participants who were readily available (Teesside University administration staff) were used to evaluate the four prototype interaction styles.



Participants were *not* selected to be samples of the intended population as this study was seeking to establish if design guidelines could be derived by quick and dirty user trial methods. Each participant selected one variant from each pair based on the following assessment criteria. 1) Which was most easily learnt? 2) Which would be the easiest icon to use during constant use? 3) Which was most preferable in terms of visual appearance on a microwave control panel? Simple accumulative scores from each criterion were used to select the favoured solution. A touch screen monitor was not available at the time so studies were conducted using a mouse rather than a touch screen. Participants were encouraged to treat icons as if controlled by a touch screen (for example reaching up and touching icons on the screen to see if feedback given was still appropriate).

First, variants on icon selection feedback were presented (see Table 4.14) for evaluation; secondly, variants on multiple-choice selection techniques (see Table 4.15); third, four icon animation styles for different frame loop sequences (see Table 4.16), and finally six different variants of icon dragging behaviours (see Table 4.17).

Ranking	Component Interaction style variants
1	Box frame appears around icon on contact
2	Black and white to colour on contact
3	Green to red colour on contact
3	Sound on contact
5	Highlight 'glow' of icon on contact

Table 4.14 Ranking of selection feedback variants

Participants commented that the 'box frame' made the most difference and 'looked like a button' (Table 4.14). Some participants also indicated that a second, additional variant could be combined to increase feedback. For this



reason, the design of this interaction style component combined the top two variants.

Ranking	Component Interaction style variants
1	Selecting an icon, a duplicate icon appears in destination box - select icon again - another duplicate icon appears adjacent to first icon but icons begin to 'jigsaw' together to form a whole object (Pizza slices/whole pizza)
2	Select icon, a duplicate icon appears in destination box - select icon again - another duplicate icon appears adjacent to first icon to form a line of objects (pizza slice used as object)
3	Selecting an icon resulting in a duplicate icon appearing in destination box - select icon again - another duplicate icon appears adjacent to first icon to form a line of objects (whole pizza used as object)
4	Select different icons each depicting different numbers of objects e.g. 1,2,3 pizza slices.

Table 4.15 Ranking of multiple selection variables

A multiple selection task was presented to participants. This was based on using an 'auto-cook' function, which altered power and cooking times depending upon the number of pizza slices selected (Table 4.15). Subjects preferred selecting objects in such a way that they were building up a picture of a complete object.

In reviewing the animated icons, participants commented, that although longer animations were more 'eye catching', they may become annoying over time (Table 4.16).



Ranking	Component Interaction style variants
1	Three frame animation looped twice
2	continuous three frame animation
3	Three frame animation looped three times
4	Three frame animation looped once

Table 4.16 Ranking of animation icon formats

Results suggested that the participants preferred the icon to move directly to the target (Table 4.17) as this was faster and required less dexterity to complete the selection. Users were also in favour of ‘ghost’ image on contact as they too gave useful icon status feedback.

Ranking	Component Interaction style variants
1	Touch icon, icon instantly moves from home to default target
2	Touch icon then touch target - icon moves from home to target
3	Touch icon (selection identified by ‘greyed-out’ or ghost image) then touch target
4	Drag icon to target, leaving ‘ghost’.
5	Touch icon then touch target
6	Drag icon to target

Table 4.17 Ranking of dragging behaviour variants

4.4.2 Final Prototype

A fully functional prototype interface was completed and composed of pre-set time settings (1 min, 1.30 min etc) the use of animation to describe selections, countdowns, de-selections, user definable settings to allow the product to adapt to individual usage patterns and more descriptive and usable ‘auto cook’ functions. Figure 4.2 illustrates the final prototype



interface. Some of the icons are non-functioning in this state and therefore 'greyed-out'.

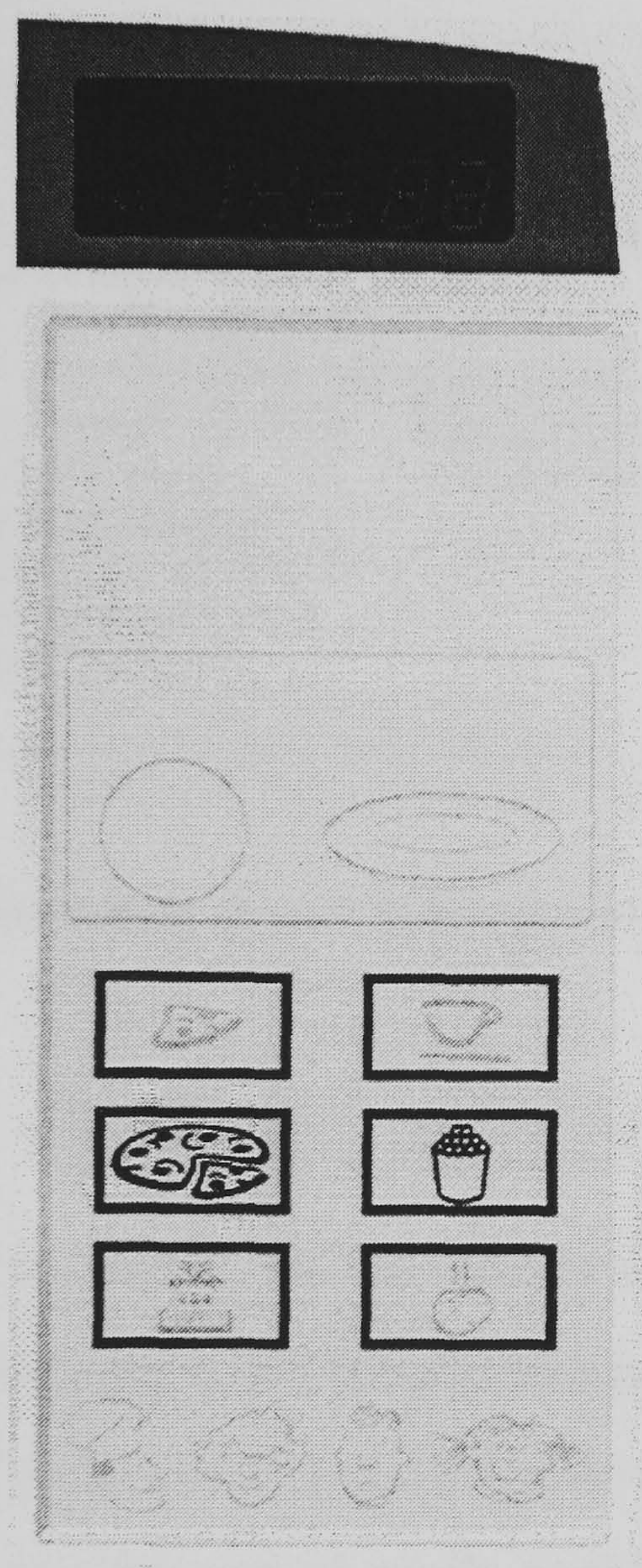


Figure 4.2 Prototype for 'adaptable' microwave interface

**4.4.3 Guidelines derived from study**

Guidelines were derived from the results of the evaluation studies and from observations made during the studies. Summary guidelines derived from the results are presented in Table 4.18 and include design process guidelines based on lessons learnt from using low and high level prototypes.



Design guidelines	Derivation
In selecting icons using a touch screen, try to create 'closure' by framing the icon once it is selected. Background or icon colour changes on selection can also improve selection recognition.	Ranking of icon selection variants
When selections are being made from a finite list, ensure the possible minimum and maximum selection options are obvious	Ranking of multiple option selection variants
Animated icons should be presented for a duration of around 1-2 seconds or be continuous	Ranking of animated icon formats
Icons should not be dragged to a target destination but move to home or default target automatically when selected	Ranking of dragging behaviour variants
For the design and evaluation of novel interaction styles, plan for spontaneous and frequent informal usability tests. Ensure that as much contextual information is offered to the participant during the trial. Always consolidate with a more formal usability evaluation using high level prototypes.	Experience gained from using high and low level prototypes

Table 4.18 Guidelines derived from quick and dirty studies

4.4.4 Discussion

As with Study 3, guidelines derived in this study were neither extensive nor groundbreaking. The reliability of these guidelines has obviously been weakened due the use of subjects who may not be entirely representative of the intended population and the use of crude ranking methods. This was thought a worthwhile compromise to gain guidelines supporting immediate usability design problems and introduced users into the decision making process. This approach allowed unknown usability issues to be quickly considered and design 'fluidity' to persist longer. Study 2, suggested that one area where the content emphasis of guidelines could be improved was at points where design 'options' or 'criteria' were considered. The design of low level prototypes provided potential options while participant preferences provided a notional set of preference criteria on which to form design judgements. Design decisions using this approach did sometimes cause conflict. For example, when preferred icons were combined such as a static icon for 'auto-defrost' with other animated icons it produced



inconsistency in the interface design. This only became apparent when static and animated icons were combined to form a single interface prototype. Results from the user trials could potentially be too fragmented and isolated to gain any useful design data.

Another difficulty in gaining good design data was the lack of contextual relevance that the participants experienced. Many participants used in this study admitted making decisions based purely on arbitrary subjective preferences and were not basing their preferences within the context of cooking tasks. Evaluating an animated icon, for example, was too isolated and removed from any tangible activity to make a real informed decision.

The use of spontaneous usability tests did, however, prompt more open debate between the interaction designer, participants and researchers about how novel interaction styles could be made more usable.

## 4.5 Conclusions

Both approaches to proactively producing CPT guidelines were not feasible due to the inherent novelty of the interaction styles produced. In Study 3 the lack of sufficient, contextually relevant usability design data, during the development of high-level prototypes, resulted in 'gross' usability problems being carried over to the evaluation. This reduced the number of detailed design guidelines that could emerge from the study findings. Study 4 attempted to overcome this by solving usability design problems spontaneously using low level prototypes with 'quick and dirty' user trials. However, participants used in these studies found it difficult to make informed judgements. So again, useful design guidelines could not be produced.

For these reasons, guidelines produced from these studies contributed little to the draft CPT guidelines evaluated in Chapter 5. Nevertheless, consolidating the findings from these two studies later helped to direct



research towards identifying an alternative bridging mechanism in Studies 7 and 8. These studies were published in a conference paper (Bonner and van Schaik 1998).



# 5 Evaluating CPT guidelines in commercial design projects

---

## 5.1 Introduction

During the literature review (Chapter 2), studies measuring the effectiveness of guidelines were criticised for being centred around contrived experimentally-based design problems or exploring very specific aspects of guideline effectiveness. Guideline evaluation studies were not found using field-based or longitudinal methods. To counter this, two evaluation studies were designed to address these gaps in the literature.

Study 5 was a precursory evaluation study carried out shortly after Study 1, where potential users (interaction designers) influenced the content of guidelines. This provided an opportunity to implement a user-centred approach to guideline development by improving guidelines produced in Study 1.

In Study 6, prototype CPT interface design guidelines were developed based on findings and emerging results from Studies 1 - 5. Three small evaluation trials were conducted on the prototype guidelines. Both studies were assessed against the effectiveness criteria developed in Section 2.3.

## 5.2 Aims and objectives

- To produce and evaluate CPT guidelines based on user-centred design principles
- To produce a draft set of CPT guidelines based on findings from Studies 1-5
- To evaluate the draft set of CPT guidelines against the effectiveness criteria



### 5.3 Study 5 - Evaluating guidelines based on designers' requests

The aim of this study stage was to produce a first generation of CPT guidelines based on requests from potential users of CPT guidelines. The aim was to improve the applicability of guidelines by allowing the context of 'live' design projects to be considered in the development process. The first version of guidelines was generated through user-centred design methods.

#### 5.3.1 Method

Four design departments in large manufacturing organisations (Electrolux Design Group, Spennymoor; Raychem User Centred Design Group, Swindon; NCR Usability Department, Scotland; Canon Advanced User Interface Design Group, Guilford) were approached by telephone and asked if they would be prepared to evaluate interface design guidelines specifically tailored to their needs. After agreeing, they were asked to draw up a set of usability design questions that would assist on a current interface design project. Each contact person was asked to initially supply no more than 6 questions, to reduce any anxiety about the task being burdensome. Further support was not provided in terms of what type or level of questions to pose merely that questions should reflect issues within a current interface design project. Designers were told that, once questions were received, design guidelines would be derived using the database or based on guidelines already developed. It was explained that, at this stage, guidelines would only be presented in plain text format to provide a catalyst for critical comment. They were then asked to use the guidelines in a way most suitable to their chosen design project. Guidelines were sent by post. The intention was not to provide any further encouragement or prompting during this period to prevent any sense of obligation or enforcement. However, some contacts did require further assistance in devising questions. Interviews were then conducted after a four-week period in which the following effectiveness criteria were considered:

- relevance to novel interactive devices
- content emphasis



- usability
- relevance to target audience.

### **5.3.2 Results**

The most interesting observation was the difficulty respondents had in producing a request list or set of questions. The contact at Canon needed further clarification on two occasions about the level of questioning that was thought to be appropriate. Despite several conversations to assist in devising questions, he eventually retired from the study. This may have been due to reluctance in appearing to be making obvious or simple requests or being unable to compose a list without direct reference to a particular design problem. Respondents certainly had difficulty in making requests out of context. Requests that were made tended to be for predictable, conventional ergonomics data such as size, shape, protrusion and forces required for operating control buttons, icon design and the use of foreground and background colours in displays.



---

## Visual displays

Should be

- legible
- read left to right
- spacing between numbers should be between 6 mm and 12 mm
- electronic displays should be free from brightness variation, jitter, flicker as possible
- black on white preferred.

The minimum acceptable display size of a display depends upon the type and amount of information presented at any one time and the viewing distance.

## Digital displays

- can be read very accurately and quickly
- not appropriate for displaying information about rates of change
- require minimal space
- avoid vertical arrays
- should not be too recessed
- spacing between numbers should be between 6 and 12 mm
- slopes on electronic displays should not exceed 11 degrees.

## Status indicators

Lights can be used to draw attention to the following functions, power on, function active, special hazard condition lights, operating mode light, malfunction light.

- use flashing lights to
  - attract further attention
  - request immediate action
  - indicate a discrepancy between the required state and the actual state of the system
  - indicate change in progress (flashing during the transition period).

In contrast to industrial and commercial products, coloured lighting displays very often do not conform to standards such as BS4099.

## Flat panel displays

Light emitting diodes (LED)

- assembled to form medium sized panels (e.g. 30 x 40 cm)
- element frequencies in these arrays may approach 100 /cm
- individual lights only omit one colour.

Plasma displays

- a large plasma display can measure 1 metre (or more) in diameter
- displays are usually between 25-50 pixels/cm
- luminance of display between 35 - 1700 cd/m<sup>2</sup>
- acceptable range of viewing angels is up to 80 degrees from the normal.

Liquid crystal displays (LCD)

- limited viewing angle, no larger than 45 degrees from normal
- tilt adjustments useful so users can compensate for any glare

LCD displays have to display small segments of text legibly. This can be done in two ways:

- leading - shifting the text from right-to-left across the display so that new information is always coming in from the right
- RSVP - (rapid, serial, visual presentations) where each display contains successive words or other text segments without the appearance of motion

When text must be presented in a small screen area such as that words must sometimes be truncated, then the best method of display is RSVP for relatively rapid presentation rates.

---

**Figure 5.1 Sample of database-derived guidelines**



However, some questions were more challenging such as, *'what tone of messaging do users of public information systems expect/prefer, e.g. friendly, chatty, personal or do they expect short concise, meaningful messages that are non-human like. Also how is this question affected with intelligent and non-intelligent devices?'* and another example, *'What do people think intelligent consumer products are?'*.

For the remaining respondents, a small set of guidelines was offered. There was some duplication of requests resulting in a relatively small number of guideline sets, these included support on input devices, touch screens, display and screen design, general control panel design and general principles on user centred design. A sample of these guidelines is provided in Figure 5.1. Interview results are presented by organisation.

### *Electrolux*

The guidelines provided to them were not used heavily during the evaluation period, as it did not coincide with a design project focusing, at that time, on interface design. Nevertheless, the designers had referred to them occasionally. The role of guidelines was discussed along with specific comments. Designers remarked that only their own product graphics guidelines were used heavily and generally felt that guidelines reduce creativity and limited solutions. If they were to be improved, they should be written in a form that suggests multiple solutions to an interface design problem and allows for interpretation and provides a contextual framework in terms of where and when they could be used.

Paradoxically, they suggested that guidelines provided positive support, primarily by adding credibility in design meetings or as a mechanism to reduce conjecture and provide a clearer decision making path. Finally, adhering to guidelines was regarded as being beneficial to the consumer by reducing 'mis-use' of the product, although there was also concern that design guidelines could be used to impose a rigid design style or approach, or as 'a stick to beat designers with, if used incorrectly'.



Specific comments about the guidelines were generally negative. One point that was unanimously made was the lack of graphical examples. Data should be in a clear and easy to read style as some of the text was considered to be too technical to be useful. Tabular information was liked and felt to describe information in a simple digestible format. Text based guidelines, regardless of their content, were considered to be 'off-putting'. Guidelines needed to be 'qualified' and state how important the information was and include the derivation of the data.

### *Raychem*

Three designers from Raychem provided feedback on the prototype guidelines. Their opinions were similar to Electrolux. Guidelines were considered useful to clarifying confused thinking and prevent design errors and could be used as a 'back-up' for design decisions. They could also be used as an 'aide-memoir' or checklist to ensure that all design parameters had been considered. They preferred to use guidelines in a more retrospective role only using them when some aspect of the design was incorrect.

Specific comments about the guidelines were, again, generally negative. Guidelines provided were often thought to be stating the obvious and perceived as imposing a formal design process by creating a more 'engineering' based approach to the design process. The guidelines were criticised for having too much text and few illustrations and lacked references to useful case studies.

Speed of access was thought important and it should be possible to dip into the guidelines quickly and be supported by cross-referencing. They suggested that ambiguity was a positive aspect as it allowed interpretation and could encourage alternative options to be explored. One designer suggested that future guidelines should provide an overview of a subject area including current theories, boundaries of interest or terms of reference;



influences and assumptions made on the data, checklists or decision trees and a list of definitions.

#### NCR

A design team including industrial designers and a psychologist used the prototype guidelines on three different design projects. On the whole they thought the guidelines were ineffective because they lacked prescriptive data and had to be interpreted to be of use. In some situations the guidelines did not cover the topic areas as originally requested, as this information was not available in the database. Checklists were suggested as a possible structure, even if some of the items on the checklist could not be substantiated with research evidence or a rationale. One designer suggested constructing checklists using a style such as '*If I were the designer/evaluator, I would consider a, b, c...*'. Referencing other guidelines within a guideline was disliked and it was suggested that each guideline should be self-supporting and independent. Tables of information worked well and should be used more often.

### 5.3.3 Discussion

The problem of question composition had not been anticipated at the beginning of the study but supported similar evidence found in Study 2 (Diary observations). Designers had difficulty in framing a design problem through explicit questions. Both studies demonstrated ambivalence towards design guidance. Guidelines should not hinder creativity but support it, yet they should also ensure design solutions are checked to conform to good design practice. No evidence was found to suggest usability defects were systematically addressed either with or without the use of guidelines (for example using ergonomics criteria, see section 2.5.3). Responses from designers suggested that interface design tasks were tackled through 'problem-encounters' where usability issues were dealt with as they emerged. Again, this was consistent with observations found in Study 2 where identified or known usability problems were resolved through personal



'experiential' knowledge and unknown or irresolvable problems were carried through to the usability study.

Although the designers offered constructive criticism, discussion centred on the structure, presentation and relevance. Few references were made to recalled situations where guidelines were actually used and most interviewees discussed general problems of using guidelines.

Conflicting and consensus statements are summarised and phrased in terms of design recommendations below and were used in Study 6.

### *Conflicting recommendations*

- Each guideline should 'stand alone' *conflicted with* support cross-referencing
- Importance or value ratings should be assigned to each guideline in order that designers can determine how to judge their importance, should suggest a range of solutions or alternatives and allow for and support interpretation *conflicted with* checklists should be provided to ensure that all design issues are considered or quick reference facilities should be provided

### *Consensus recommendations*

- Examples should be frequently given in a variety of forms such as case studies and diagrams
- Should be divided between overview and detailed material
- Should be 'visually' interesting and not rely heavily on text to convey information



## 5.4 Development of guideline design requirements

Guideline design requirements identified in Studies 1 - 5 were drawn together for a more comprehensive evaluation study (Study 6) and were structured around the eight *specificity* and *applicability* criteria. These design requirements helped shape and define a sample set of prototype CPT guidelines (Appendix 13.4).

Study 2 suggested that content emphasis should be geared towards design process activity, Study 5 also indicated a need to support both divergent and convergent activity by allowing a design problem to be checked for compliance to good usability practice. Observations in Study 2 suggested guidelines should supplement designers' preference for experiential knowledge. There was also clear evidence from the literature review, Studies 2, and 5 that guidelines were not used when viewed as constraining or requiring interpretation.

These were difficult requirements to contend with. The content emphasis of the sample guidelines was presented in alternative ways in an attempt to accommodate these requirements. Guidelines were offered, as far as possible, as design rules but different ways of considering these rules were offered, such as:



- offering related design examples

Here are some metaphors to describe collections or categories of component objects, some of which also provide natural relationships between components.

Metaphor	Components of metaphor model
Desk	Drawers, files, folders, papers, note cards
Publication	Books, newspapers, newsletters, articles, figures, forms
Television	Programs, channels, networks, commercials
Music	Tracks, CD, Tapes, Music charts,
Games	Boards, cards, game pieces
Film	Rolls of film, slide holders, still and moving images
Storage	Shelves, boxes, filing cabinets, folders, paper clips
Trees	Branches, trunk, leaves, roots

Here are some examples of action metaphors, these are often more difficult to select and describe.

Action	Types of mental model
Browsing	Window shopping, thumbing through books
Selecting	Touch items, place boundaries around items
Deleting	Putting items in a waste bin, rub-out
Assign values	rotate knobs, move sliders, select items over time

Figure 5.2 Example of guidelines offered at design options



- introducing checklists, this had been considered a popular approach to presenting data

However the best approach is to understand, as fully as possible, the tasks that will be undertaken. To do this, follow these steps.

- ☐ Do a task analysis, in this context, this involves the study of user activity achieving a particular goal ⇒009.
- ☐ Identify critical steps and procedures
- ☐ Design to accommodate critical task elements
- ☐ A user trial should usually be conducted in order to find out the answers to the following checklist ⇒ 008. This checklist will help check if normal tasks patterns have been considered.

	✓	✗	N/A
Has a user profile been drawn up that considers the range of users that might use the product? ⇒ 003			
Has all the possible types of tasks and environments (untypical as well as typical) been listed and considered? ⇒ 007			
Is the interface 'transparent'? - allowing the user to focus on the task and not be hindered by the interface ⇒ 001			
Do users always know what to do and how to accomplish their task?			
Does the interface fit with user expectations? ⇒002 ⇒010			
Does the interface allow the user to explore and make mistakes? ⇒ 012			

Figure 5.3 Example of guidelines offered as a checklist



- using 'things to think about' sections, providing alternative or controversial statements for consideration

---

### Things to think about

Some products such as washing machines require anti-tamper devices to prevent accidental or mischievous use by children. Whilst some products are designed specifically for their use, there is not a significant body of knowledge about characteristics and abilities of children which can be translated into design criteria. Some development work on a communication tool, however, called PenPal suggests that children prefer products with physically manipulable interfaces and a 'fuzzy boundary' between the physical and software elements of the product. Products should be fun to use and it should be possible to customise the interface.

Products have traditionally been designed for use by people during 'active adulthood' although there is increasing recognition of the elderly user. Elderly users require a different type of interaction support. Some research investigating an elderly user group and their opinions on consumer products and found that they liked using instructions to learn how to use the product although they complained about the clarity of the instruction material. It was also found that the size, number and spacing of the buttons greatly affected ease of use.

For most disabled users, it is the variety of controls found on the product that causes concern. Furthermore, if one key control is inoperable by a user, the product may be unusable for that person. A preferred method for designing for the disabled is to use less variety in the selection of controls on each product.

Other factors that determine differences in aptitude are: experience both personal and culturally; 'technical aptitude' defined as spatial and reasoning aptitudes and task specific knowledge. However, it has been argued that other personality based characteristics are weak and inconsistent predictors of performance.

---

**Figure 5.4 Example of guidelines offered as discussion and comment**



- presenting the advantages and disadvantages of different design approaches or solving usability problems

The use of metaphor does have to be used with some degree of care. The only way to find out if they are successful is through some form of user trial ⇨008.

Advantages	Disadvantages
Can convey functionality and relationships quickly to the user	Interpretation of metaphor may be taken too literally
Easy to learn and remember	A mixture of metaphors may build in contradictory cues
Can simplify the interface	Limits of metaphor may be difficult to discern
Invites exploratory learning	Mapping between the product functionality and the metaphor may be too weak to understand
Reduces anxiety	Difficult to describe abstract concepts such as pre wash or reheat

Figure 5.5 Example of a guideline providing evidence for the use of metaphors

Accessibility to guidelines, a common problem found in the literature review and Studies 2 and 5, was improved by providing cross-referencing between guidelines, indexing, guideline checklists mapping the design process while also offering more conventional prescriptive, device-based guidelines.

Guidelines were designed with strong visual cues, navigational aids, tables, illustrations and examples. Strong evidence emerged from Study 5 on the importance of visual presentation. It was hoped that this would help infer design imagery and concepts. The prototype guidelines were not complete but a sample of potential guidelines. They did not include all aspects of



'ergonomics criteria' although many of the criteria were implicitly included in the sample guidelines.

Studies 3 and 4 had revealed how increasing the specificity of new guidelines to novel interaction styles was difficult. An attempt was made to resolve this by providing guidelines that contained suggestions for mini-experiments to alter the way design options and criteria are considered. To reduce design decisions being based on assumptions about user behaviour, guidance on user behaviour and acceptance, such as conforming to stereotypical behaviour and considering different user types, was provided. The intention was that designers could then reflect on this information thus enabling them to make more considered judgements.

---

Related Guidelines	002 005	
Contradictory Guidelines	003 004	
Rating	****	*
	Novel interfaces	Simple or commonly used interfaces

---

Figure 5.6 Example of guideline rating scale

Attempts were made to improve the validity of the prototype guidelines (interpretation of the meaning and significance of a guideline) by providing a rating scale for each guideline. Rating values were applied indicating the appropriateness of applying a guideline within a particular context. A high star rating meant the guideline should be taken very seriously, while a low star rating suggested a context in which the guideline had less importance. Contradictory and related guidelines were also highlighted if appropriate (using reference number to other guidelines), to increase their contextual



validity. Improving the reliability of the guidelines (consistency in resolving the same type and number of interaction design issues across different design problems) could not be addressed until the prototype guidelines were evaluated.

Table 5.2 summarises the key design recommendations that have been drawn from the research studies to-date.

Specificity Criteria	Applicability Criteria
Content emphasis - Designer-centred, supports convergent and divergent thinking through different problem solving strategies. Support mini experiments; provide support on behavioural models of the user.	Usability of guidelines - Guidelines made more contextually relevant by offering examples, checklists, improve accessibility by offering a range of searching features
Structure and presentation - Use of strong visual cues, navigational aids, tables, illustrations and examples	Reliability and validity of guidelines - Rating scales were implemented to improve validity
Relevance of guidelines to 'ergonomics criteria' - Guidelines produced were a sample set, so not explicitly considered	Relevance to proposed target audience - This criterion was tested in Study 6
Relevance of guidelines to novel interactive devices - Simple behavioural models presented to allow more informed design judgements to be made during important decision making points	Likelihood of survival or usage in organisational or design context - This criterion was tested in Study 6

Table 5.1 Summary of design recommendations for prototype CPT guidelines

5.5 Study 6 - Evaluating the effectiveness of prototype CPT guidelines

The aim of this study was to establish if the prototype CPT guidelines (Appendix 13.4) were more effective having implemented the design recommendations. Two of the three evaluation methods used field based methods as concern had already been raised about the validity of measuring effectiveness in laboratory conditions (see Section 2.6). The laboratory-controlled study therefore only reviewed specificity criteria.

5.5.1 Methodology and results

Three different evaluation methods were adopted to ensure that all effectiveness criteria were considered, these were:



- Controlled student design project where students were provided with a design brief and asked to use the prototype guidelines to evaluate specificity criteria.
- Questionnaire survey with experienced and practising designers and human factors specialists, evaluating both specificity and applicability criteria
- Interviews with industrial designers at Electrolux after a two month period of using prototype guidelines to evaluate applicability criteria

### *Student project study*

Twelve BA (Hons) second year industrial design students at Loughborough University volunteered to undertake a microwave interface design project. Students were asked to give particular attention to the development of alternative interaction styles that would increase the usability of the interface. The project was carried out outside of their normal curriculum activities. An introduction to the project was given with technical information on the microwave provided. They were instructed that interface design should support interaction tasks based on a task flow chart given to them, which described functions and action sequences. Students had a four-week period to complete the project and were asked to approach design activity in a similar manner to other design projects but address any design problems by using the prototype guidelines. Once completed, they were asked to answer a questionnaire. Semi-structured interviews based on specificity issues were then conducted with students who had completed the project. No incentive was given other than their work would be shown to design staff at Electrolux for possible temporary employment.

Only four students submitted their design work and questionnaire. These students were interviewed and were asked to elaborate and discuss their views on the guidelines. Two other students submitted their work later by



post; one of which had completed the questionnaire and the other provided written feedback. Responses were very positive. Results from the questionnaires revealed that examples, illustrations and checklists were popular. There was concern over some guidelines having too much detail, particularly the descriptions of mental models. Navigational aids were thought to be too 'heavy-handed'. During the interviews, students added that the guidelines helped them to think more methodically and ensured that all usability issues were addressed. The students liked the structure the guidelines imposed in addressing interaction design issues. Most commented on how the guidelines offered them information previously unknown to them in an accessible manner. A review of their submitted design work indicated that they had taken much of the design guidance very literally with some guideline examples inserted directly into their design proposals.

#### *Questionnaire to practising designers*

Telephone contact was made with 25 respondents known to be interested in interaction design issues who were also willing to carry out the evaluation study, of these 20 were designers and 5 were human factors specialists. Respondents were sent a copy of the prototype guidelines along with a questionnaire and asked to use the guidelines as extensively as possible or to appraise them if not used during the evaluation period. After four weeks, participants were telephoned and reminded to complete the questionnaire if they had not already done so. Guidelines were distributed to HF specialists to evaluate the guidelines as 'content specialists' and were not expected to use the guidelines to support design activity but as users of HF guidelines.

Six designers and three human factors specialists responded. Only four of the six designers returned questionnaires, the other two responded with written feedback. Responses from the questionnaires were all very favourable with most responses gaining high 'useful' scores. Most of the criticism was levelled at the use of the 'rating scores' (see Figure 5.6) which were found to



be didactic and navigational graphics which were found to be confusing and distracting. One respondent suggested that while the guidelines were packaged in a useful manner he did not feel that guidelines would alter design behaviour. The two human factors specialists concurred mainly with the designers. One of them expressed concern at how designers might interpret some of the psychological descriptions posed within the guidelines.

#### *Use of guidelines in a commercial design studio*

The prototype guidelines were given to four industrial designers working in the Electrolux Design Group. They were given a brief introduction to the structure and content of the guidelines and asked to use them during the following two-month period. No incentives or inducements were provided. At the end of this period, participants completed the questionnaire and were interviewed.

The use and uptake of the prototype guidelines was extremely poor. During the introduction of the guidelines, the designers gave favourable responses and thought the content and presentation would be useful. However, in practice the guidelines were not used. During interviews with the designers they commented on forgetting to use them, as they were not an integrated component of their design management process. They added that display and control interface devices were 'sourced' components; for example, encoders and displays are manufactured by other organisations thus limiting interface design freedom.

#### **5.5.2 Discussion and consideration of more recent research**

The evaluation studies revealed that specificity had been improved. Many positive comments were found on how the prototype guidelines were an improvement on other existing guidelines in terms of content and visual appearance. Most designers were very positive, particularly the student designers. This may be due in part to their design inexperience in using guidelines or willingness to follow an assignment-based study.



Nevertheless, designers began with very favourable opinions on initial contact with the prototype guideline (perceived usability of the guidelines). These views did not follow through with their subsequent usage or implementation (survival and relevance). Although designers responded positively to the guidelines, this did not guarantee successful uptake or influence interface design practice. Even though all three studies used small sample groups and the response to the evaluation studies was weak, this evidence was consistent across all studies. Designers can clearly articulate what they need from guidelines and provide useful feedback to improve them but, despite this, do not use them in practice.

During the latter stages of the guideline development studies a set of interface design guidelines for consumer products was reviewed known as the 'TRON' guidelines (Sakamura 1993). This document was not in the public domain and therefore not identified in the literature review. The TRON guidelines presented interface design guidelines on control and display devices more commonly found on contemporary products and machines. Some examples included, time setting devices, multi-function controls, batch processing (setting of a range of parameters before selecting 'start' button), design and manipulation of 'settings', and field and display scrolling. One of the interesting aspects of the guidelines was how device components are described using a generic or product free taxonomic notation in order to distinguish between different hardware and software components. The guidelines can only be used once this nomenclature is understood. It is unfortunate that awareness of these guidelines came too late to introduce them into this guideline evaluation study. However, it was unlikely that they would have had any direct impact on the final direction of this research due to the reluctance of designers to use any types of passive design data no matter how convincing its specificity.

The ineffectiveness of guidelines was reinforced by another important study that had been recently published at this time. As had already been reported



in the literature review, studies had not been found that evaluated the effectiveness of design guidelines using both issues of specificity *and*, more importantly, applicability. However, Burns *et al* (1997) undertook a study (published after the literature review was undertaken) examining the effectiveness of guidelines within a more 'naturally' applied context. Three effectiveness criteria were independently evaluated, usability, usefulness and viability in three studies using human factors handbooks (which were not specifically written as guidelines). The three criteria offered similar coverage but different emphasis to the criteria used in this research.

In the first study they attempted to identify factors that may influence ease of use by comparing two human factors handbooks which used different data structures. The text with more applied rather than theoretical information performed better, although, consistent with the studies here, subjects found the data in both books to be too general and lacking contextual information. In the second experiment, usefulness of design data was evaluated. A small sample of human factors specialists in the nuclear power industry were asked to assess the usefulness of a human factors text. Subjects were asked to find answers to a set of questions and rate the answers in terms of cost; importance of the answer to questions provided to them, relevance and effort. One of the key findings was that 'professional HF designers prefer to miss information rather than waste their valuable time looking in handbooks' (p 318). The final study examined viability, which identified factors that might influence whether HF handbooks actually get used. This was a laboratory study using HF students rather than field-based. They found that students did not regularly refer to the human factors text during their design work, again this is consistent with findings in Studies 2 and 6.

They stated that human factors design data will only be used if it is highly application and context specific, that human factors data needs to be industry specific and suited to designer's information search behaviours, and commented that designers are conservative searchers. Indeed, findings from



their research suggested that even if design data is application and context specific, it will still have minimal impact in terms of effectiveness. They conclude, 'that handbooks are not the final answer to the problems that plague our discipline. Even if one had the perfect human factors handbook, it would probably only have a low impact on design. Consequently, we, as a discipline, need to search for other ways of bridging the gap between what we know and what we do' (p 321).

## 5.6 Conclusions

As described earlier in the literature review, Rouse and Boff (1998) stated that few studies had addressed how ergonomics data and information should be presented to designers. Study 6 concluded a series of related studies in this research attempting to do this. However, the findings from these six studies concluded that guidelines, even when consciously developed and presented for designers, still have been shown to be ineffective if presented using a passive medium. The Burns *et al* (1997) study concurred with these research findings.

These combined results make an important case against the use of guidelines as a bridging mechanism between research and practice. Human factors, interaction design or HCI researchers should not provide design data as a passive repository. Improving the specificity and applicability of this form of design data had limited impact on improving effectiveness. Publishing or reporting findings from usability studies would appear to be a very ineffective approach to improving the usability of user interfaces. The need for a common representational framework between different design disciplines as argued by Rouse *et al* (1991) still remains. The research question posed in Section 2.8 remained unanswered although the inquiry so far provided an indication of how design data should not be presented and where possible solutions may lie.



Other bridging mechanisms needed to be identified. Further studies were introduced focussing on the organisational context in which interface design activities occur by exploring applicability factors using a broader frame of reference.



## 6 Exploring alternatives to guidelines

---

### 6.1 Introduction

Conclusions drawn from the previous studies suggested that interface design guidelines are unlikely to be used by interaction designers. Alternative bridging mechanisms between research and practice needed to be sought. Research was broadened to consider the organisational context in which ergonomics, HCI or interaction design data was used. Study 7 reports on a telephone survey exploring the use and acceptance of ergonomics methods in commercial UK industrial design organisations. The second study, Study 8, made comparative observations of two design organisations that used design data in radically different ways. These two studies were written as a refereed conference paper and presented to the 'Design Cultures' conference, Sheffield Hallam University (Bonner 1999b).

### 6.2 Aims and objectives

- To understand how extensively research-based design data are used in commercial industrial design organisations
- To compare conventional and more radical methods of developing interaction design data

### 6.3 Study 7 - The use of research-based design data in UK design practices

The purpose of this study was to establish what, if any, type of ergonomics, HCI or interaction design data are currently used in industrial design practices. The term 'ergonomics methods' was used in the study to collectively describe any form of interaction or user centred consideration and because it is a commonly known term amongst industrial designers. Questions were posed to elicit examples where they had been successfully implemented into the design process. In this way methods with high



applicability could be identified. A telephone survey was conducted which covered a high number of design organisations quickly and cost effectively. Large sampling was possible, as the number of design consultancies in the UK was relatively small.

### **6.3.1 Method**

A list of 122 practising design consultancies in the UK was obtained from the 'Directory of Designers' (1991). The survey was designed to take no longer than 10 minutes. This was deemed a necessary constraint in order to gain more willingness from the respondents to take part, but limited the number of survey questions that could be asked. A questionnaire was drawn up covering the following topic areas:

- General information including name of contact, position in organisation and number of designers employed
- Type of design work undertaken, distribution and involvement of work in types of design activity
- Involvement of users in the design process (if applicable)
- Ergonomics methods used, type of interaction or interface design methods used
- How, if appropriate, they would like to see interaction design data be implemented in the future

Questions were read out to each respondent and opportunities were provided for the respondent to expand upon a point if necessary. A small pilot study was conducted with two organisations. A less formal and unstructured approach was required as respondents tended to raise information not related to questions asked or pre-empted other questions.

A research assistant contacted design organisations, where a senior designer or design manager was sought. Once contact had been made, a brief explanation of the purpose of the survey was given. If the respondent



agreed, the interview was conducted or, if necessary, an appointment at a more convenient time was made.

6.3.2 Results

Attempts were made to contact all 122 organisations, which had identified product or industrial design as part of their expertise. From this directory list, 21 had closed, 9 declined to take part in the survey, 15 organisations were rejected due to the difficulty of contacting a relevant individual, and 27 stated they undertook little product design work. Therefore 50 organisations took part in the survey, of these, 2 were used in the pilot study. Of those contacted, 27% of the respondents described themselves as either the proprietor or director of the organisation, 64% as practising designers and the remaining 9% had other roles.

Response	Frequency
	(n = 48)
Other	28
Consumer products	25
Industrial	12
Medical	11
Graphic design	4
Computing	4
Transport	3
Telecoms	3

Table 6.1 Response frequencies for design specialisms

The mean number of employees in the organisations surveyed was 18.5 ranging from a minimum of 1 to a maximum of 60. The mean number of designers employed was 5.2 within a range of 1 and 22. Mode and median data can not be provided as a few respondents returned range or fluctuation values for employees or designers stating that they vary depending upon the



project. In many situations freelance or consultant personnel maybe temporarily recruited. Mean values were calculated by taking the middle value of any range provided.

Respondents were asked what areas of industrial design work they specialised in (Table 6.1). Most organisations tended to specialise in 2 -3 different areas. The results indicate that 25 (52%) considered 'consumer products' as part of their portfolio. Many organisations identified a wide variety of design activities, classified as 'other', which included packaging, tablewear, furniture and sports equipment.

Respondents were then asked if they had any involvement in product control or display design, 37 (77%) replied they had some involvement. They were then asked what stages in the design process they would be able to influence interface design.

Response	Frequency (n = 37)
Providing design solutions	33
Developing prototypes	30
Evaluation of solutions	26
Establishing user requirements	18
Other involvement	7

Table 6.2 Type of interface design involvement

In Table 6.2 it can be seen that 48% of respondents conduct some form of user requirements analysis and 70% undertake evaluation study. Those recorded as 'other involvement' included design activity such as graphic design or integrating corporate identity. More specifically, respondents were asked



what their design role would typically be in interface design. Table 6.3 presents the generalised responses.

Response	Frequency (n = 37)
Design elements not done by client	9
Design of interactive elements, not internal mechanisms	8
Design of all aspects of controls and displays	7
Design interface graphics	3
Client leads interface design process totally	2
Placement of controls	1

Table 6.3 Design role in interface design

Only 7 organisations (18%) reported any real level of autonomy over interface design although many respondents qualified their responses by stating that there was a great deal of variation between projects and between clients demands.

Respondents were then asked if they involved users in the design process, 33 organisations (69%) responded. Those that commented that users were not involved stated that clients usually possessed some form of user requirements knowledge typically through market research. Again, respondents suggested that the design process varied from project to project. Many respondents said that they would introduce user testing if the client had not already done so.



Response	Frequency (n = 33)
Observation of users	10
Interviews with users	9
Focus groups	7
In house trials	7
Varies with project	5
Analysis of sales	3
Hire consultant	3
User testing	2
'Hall' testing	1
Market Research	1

Table 6.4 Response frequencies for gaining 'user information'

Table 6.4 provides the response frequencies for gaining information about users, When asked more specifically if users were involved in the design process, 10 (21%) of organisations stated they conduct some form of user testing on proposed design solutions (Table 6.5). 'In house feedback' was defined as consultation with other designers within the organisation.

Response	Frequency (n = 10)
User trials involving design proposals	10
User feedback on client and competitor products	10
Observation of users with existing products	7
'In house' feedback	3

Table 6.5 Response frequencies for user testing



When asked what types of paper based ergonomics resources were used, anthropometric based information was most frequently mentioned (Table 6.6). Respondents were then specifically asked if they used ergonomics design standards or guidelines. Six organisations stated that they may use British Standards and two organisations use Design Council design guidelines.

Response	Frequency (n = 37)
Humanscale (Dreyfuss)	11
Anthropometric tables	10
Guidelines or Standards	8
Manufacturer's or client details	4
Anthropometric databases	2
Magazines	2
Research papers	1
Various range of methods	1

**Table 6.6 Response frequencies for paper based ergonomics resources**

When asked if they felt that other ergonomics methods contributed towards the design of product interfaces, 37 organisations (79%), felt that ergonomics methods often contributed. However, 8 of these, described computer aided design tools (CAD) as an important part of the ergonomics design process. Eight organisations stated that they would or had used ergonomics specialists particularly if they felt their 'in house' expertise was not adequate.

Finally respondents were asked, 'How could ergonomics based information and data be improved?' Very few respondents answered this directly, those that did suggested that more use could be made of the computer-based medium with more interactive data resources such as CD ROMs. Most of the respondents perceived that ergonomics should be an integral part of a CAD



package providing support and data when appropriate, particularly physical three-dimensional user models.

### 6.3.3 *Discussion*

The sample size of 48 represented a high proportion of the estimated industrial design consultancies in the UK. Attempts were made to use a more recent directory but there had been no subsequent editions of the Designer Directory, at that time. Enquires were made to establish how many current practising industrial design consultancies exist in the UK from the Design Council and the Chartered Society of Designers (CSD).

Unfortunately, both organisations were unable to provide up-to-date information. This information was not available for two reasons. Firstly, defining an industrial design consultancy was difficult as many organisations provide services across a number of design disciplines. Secondly, there were a high number of fluctuating small consultancies or sole traders. However, estimates suggest that there were approximately 120 organisations.

Although the sample is large, caution should still be taken with the findings. Only one individual within each organisation was spoken to and only for a short period of time. The survey provided little opportunity to discuss issues in detail. In particular, the variation in design approach depending on the design project was consistently stated and therefore generalisations were often offered.

Although a high number of respondents undertook interface design, there was evidence that their clients determine the functional specification of the interface, leaving designers to have either residual or peripheral input. The number of organisations who gained information about users in the design process (Table 6.4) and involve users in the design process (Table 6.5) appeared encouraging. Comments on user involvement were interpreted broadly. These questions were posed to reveal the level of user involvement and if user-centred testing was conducted.



Industrial designers used a narrow sub-set of available ergonomics data and methods (Table 6.6) with anthropometric based data providing the common resource. Interestingly, Humanscale (Diffrient *et al*, 1985), or 'Dreyfuss' as it was usually referred to, was still an important resource despite its age. This may be due to its compactness and a strong visual presentation of information. British Standards were also reported as a resource but it was not determined if these were consulted purely as physical ergonomics or HCI standards. Very few respondents alluded to the use of any ergonomics or HCI design or evaluation methods and perceived these to be an integral part of CAD software. It would have been interesting to investigate further the reasons for such a poor uptake of interface design or evaluation methods.

When asked in general terms what type of ergonomics methods might be used, few recalled any particular ergonomics methods. However, when asked more specifically about ergonomics methods, the respondents did provide evidence. This led to some anomalous findings particularly between the generic and specific questioning. Very few designers presented any suggestions related to how ergonomics data or methods could be improved. Nearly all the suggestions focused on providing computer-based resources particularly linked to CAD packages.

By placing these findings within the context of the previous research studies, the non-use of design guidelines was further reinforced. More importantly, the study revealed that very few ergonomics design methods were used, although there was evidence that user testing played an important role. Industrial designers had a strong awareness of the importance of user involvement but not necessarily a strong commitment to user-centred design due to having little autonomy over interface design. Client requirements imposed strong constraints on interface design decisions with user testing often used as a market-testing tool. Suggestions that ergonomics methods should be integrated into CAD software were interesting, but in light of



findings from previous studies, it would be unlikely that they would be used if presented in a passive medium.

Few alternative bridging mechanisms appeared to be available, which would fulfil applicability criteria. Designers did not have a strong awareness of the variety of ergonomics methods available and were limited in design freedom by client needs. However, one area that was considered for exploitation was the point at where user testing was conducted. The brevity of the telephone interview certainly constrained investigation in this area. On balance however, it was probably correct to restrict the interview to ten minutes, as many respondents were initially reluctant to take part and could only be convinced on the basis the interview would be short. The role of user testing was, therefore, explored in more depth in the following study.

#### **6.4 Study 8 - Comparing two design cultures**

This study addressed some of the deficiencies identified in Study 7, such as investigation brevity and lack of contextual relevance in questioning, while also exploring if user testing could be exploited as a bridging mechanism. Two organisations, with very different design cultures, were interviewed in depth, particularly in terms of user involvement.

Two design teams were selected, Electrolux and Danfoss. Designers at Electrolux had been interviewed before as part of investigation into the effectiveness of guidelines and were known to provide a traditional industrial design service. Danfoss, however, was identified as using quite radically different design methods (Buur, 1997) particularly in terms of user involvement. These two different design teams provided a useful opportunity to compare and contrast user-centred design management.

Danfoss is a leading producer of precision mechanical and electronic components and of intelligent 'mechatronic' devices, developing a wide range of programmers and time switches for automatic controls for process



control and regulation for residential and non residential buildings.

Interviews and observations were conducted at the Man Machine Interaction group based in Denmark. During the interview period the group was exploring the possibility of using TV-type remote controls on some of their products. They were also examining the use of 'smart' service tools for client organisations. The structure and organisation of Electrolux have been described in Section 3.4.1

#### **6.4.1 Method**

Interviews were structured around the themes used in Study 7, addressing perceptions of the designers' involvement in the product development process, where design projects were derived from, how end-users were involved in interface testing, and what type of design tools were used. This final area of questioning was included to explore if interaction design data could be 'piggy-backed' onto any existing design management procedures. Interviews were informal and were conducted at times when the designers were not busy. Interviews at Electrolux were conducted over a one-day period, while Danfoss interviews were carried out over two days. Notes were taken during all the interviews and interviewees were encouraged to illustrate or provide examples to support their answers such as presenting video clips, current project work or report documentation.

#### **6.4.2 Electrolux findings**

The two interviewees at Electrolux were industrial design graduates, (out of a team of four) and interviewed together. One had been with the organisation for about 5 years while the other for 18 months; they often worked on projects together and described the design team as an internal consultancy, negotiating design projects with different parts of the organisation. The interviewees were very keen to express that their design team was anxious to be involved in 'conceptual' design work where the fundamental concepts of product functionality, form and identity were



reviewed. Currently a design group in Sweden carried out most of this work.

Design skills required were 'traditional', using 2D and 3D illustration and modelling software tools with marker pen illustrations and foam modelling. Contact with users was recognised as being rare. They thought this was due the organisation's understanding of the designers' role in the product development process. Typically the marketing department gave user requirements to them.

### *Use of design tools*

Traditional design tools were used such as sketching, 3D modelling either in card or foam, 2D and 3D computer based modelling, also some style guidelines are used for different brand identities. The design group recently had to comply with the Integrated Product Development (IPD) process initiated by Electrolux. Before the implementation of IPD, one of the designers had produced an internal method, which divided into three phases of design management activity. This included:

- 'reality design' involving 'on-the-ground' or 'day-to-day' design projects;
- 'selection activity' involving the use of 'off-the-shelf' but unused technology either within or external to Electrolux, providing more innovative or progressive design solutions
- 'vision or foresight design' involving more speculative product proposals that may be manufactured in 10 years or so.

This approach was adopted for a design product but implemented in an incremental fashion rather than working on all phases concurrently. Each phase of design solutions had to be 'sold' to the group to continue funding. Acceptance for the full range of ideas could only be gained if implemented in a phased process. User studies for the third phase of the programme relied on a series of focus groups and reports provided by a forecasting



consultancy. This involved groups of potential users placing themselves in predicted 'scenario lifestyles'. These scenarios were quite detailed with each group member being assigned a name, role and status within the scenario group. From these workshops a series of design solutions were proposed.

### *Involvement in product development process*

Both stated that design activity and involvement depended upon the type of product and the type of brief given. Electrolux own many 'brand' manufacturers; in the past these were independent and still retain proprietary product development practices. This resulted in varied management structures, decision making processes and attitudes towards industrial design. Typically, most design projects were 'facelifts' where an existing design is modified mainly in visual appearance. Degrees of freedom tend to be narrow with a high emphasis on low cost or re-investment using existing or 'sourced' technology. Areas where alterations could be made tended to be in graphics, although some relocation of controls and consideration of size and shape of controls was required. The design group was beginning to receive more 'conceptual' development work with more creative design flexibility. Design solutions were printed and pinned on the wall; discussions then took place between the designers about the merits and shortfalls of each solution.

### *Sources of design projects*

A design project would be commissioned from external marketing consultancy reports, competitor analysis or through 'trend mapping' where a particular design style or market opportunity had been identified. Design projects were usually initiated by bringing together different features from a range of design proposals. Final decisions on design proposals were made by external groups, usually marketing management, or through product planning groups or steering committees.



*End user involvement*

End users were not involved during product or interface development. If ergonomics criteria were considered, then generally they were assessed on flow and sequence operations. User needs could be considered using a list entitled 'factors of influence', a checklist of psycho-social factors that affect product perception. Users were only consulted during 'clinic research' and undertaken by marketing divisions with existing or 'near-to-market' products. Products were assessed against other competitor products but purely on visual impressions; users did not interact with the products. The designers regarded findings from these studies as meaningless as they were usually presented as descriptive statistics and they found them difficult to interpret. Designers had requested involvement in this process, but were not considered necessary. The designers were very willing to involve users in design activity but thought management approval would be difficult.

**6.4.3 Danfoss findings**

Danfoss had adopted a radical and extremely innovative approach to product development. The team had been influenced by Bødker (1991) in which she argues that design methods and new user interfaces must be rooted in the specific design situation, which encompass future use situations. Interviews were carried out with one industrial designer, one computer scientist, two mechatronics engineers (one was the Man Machine Interface (MMI) group leader), a usability engineer and an 'Industrial' PhD student. Interviews were conducted on a one-to-one basis. Only one of the interviewees had any formal education in industrial design.

*Use of design tools*

Interviewees defined the process as 'empathy driven' and stated that their approach moved beyond participative design methods. Design activity usually began with fieldwork before undertaking a series of workshops involving participants and designers. In these workshops, a whole series of design tools were used and participants defined their own tasks and design



problem thus allowing them to become empowered to design products for themselves. Important milestones in the design workshops, identified either by the participants or the designers, helped define the product requirements.

A series of workshops were used to incrementally improve the interaction design problem and were conducted closely together (within a two-week period). Much of the design team's work centred on planning and analysing findings from field work and workshops. Significant or important design activity was conducted within the workshops with participants taking the more proactive role in design decisions. The process of substantiating design proposals and design detailing was conducted between workshops. The group used a wide range of evolving design tools both in field studies and in the workshops and included:

- *user characters* - allowed designers and users to consider a range of users in some form of context by providing a user profile where the character was given a name, a profession and some type of opinion or attitude towards technology
- *mapping* - users described a task or process by placing 'post-it' notes on a board or table, this helped to consolidate user group descriptions and allowed a visual map of the process or task to be produced
- *drama in design* - users went through scenarios by acting them out sometimes in front of other users where the audience act as 'drama directors'
- *scenarios* - these were used by designers to check their understanding of user's descriptions, such as conveying abstract concepts. Scenarios were developed using different approaches such as cartoons, fairy tales and a magic 'crystal globe'
- *mock-up and prototypes* - conventional foam models were used in real situations as a discussion point to evolve a design proposal



Documentation of design activity and outcomes was usually video-based. When it was suggested if other documentation methods could be used, some were wary of formalising design proposals, stating that design solutions should be 'embedded' in the workshops. They suggested that formalising evidence on paper could force subtle and complex user requirements to be misinterpreted. Projects were not usually written-up formally. One interviewee stated that one of the problems with this approach was controlling the granularity of the design detail and not knowing how viable a design solution could be. The emergence of unexpected solutions was regarded as both a positive and negative outcome of this process.

#### *Involvement in product development process*

Most of the interviewees regarded their role as a facilitator rather than designers. They placed a high emphasis on developing and understanding how they could communicate design problems and possibilities to participants, for example by trying to find new ways of presenting a product specification. Planning for a workshop was an important but time consuming task. Many of the team described how workshop procedures had changed and altered over time and undergone many iterative developments. The tools had evolved and changed to improve their effectiveness in empowering users.

At the time of the interview, two of the team had responsibility for the fieldwork of a new project. To gain an understanding of the project (a control device for a supermarket refrigeration system), visits were made to a maintenance engineer at a local supermarket, who had been used before for such design activities. Similar to the workshops, the designers engaged in planning the interview and devising scenarios in which the engineer could express his thoughts.

The team leader differentiated between the internal design process within the group and external product development activity within Danfoss. He



admitted that their approach was sometimes treated with scepticism in Danfoss, but this was reducing over time. The support of senior management was important for the survival of the team. Projects tended to be conceptual in nature with some mainstream design work. The team leader saw their future being dependent upon more mainstream or conventional design projects.

All the interviewees expressed enthusiasm for using this design approach, although some found it difficult to articulate exact procedures and define clear workshops objectives and tangible outcomes. When this issue was raised with the team leader, he explained that implicit and tacit knowledge gained through the workshops was important and many design and participant requirements were difficult to quantify. This was seen as a strength of the process.

Workshop participants were not interviewed in this study but video evidence suggested they enjoyed the workshops and became very involved. Some participants appeared more at ease than others in providing creative or innovative comments and feedback.

### *Sources of design projects*

The team leader negotiated with other Danfoss departments to gain design projects. He ensured it was conditional on their user participative methods being adopted. At the time of the interview the MMI group were also beginning to discuss design work with external organisations.

### *End user involvement*

The whole design ethos placed a strong and important role on the end-user. Participants entirely defined design activity and direction. A great deal of emphasis was placed on making participants feel at ease in adopting their design role. Video examples were shown of participants making posters, drawing and acting in order to articulate their thoughts. Designers described



some of the 'rules of engagement' which had evolved and lessons learnt to gain contextually rich feedback such as participants should only be asked to act out scenarios they had already experienced. Changes could be made within a particular scenario only if participants had had personal experience of the scenario. In one workshop they were exploring the use of portable support systems and used a scenario that all participants had experienced, in the example shown it was a device to be used during a car breakdown.

#### **6.4.4 Comparing the two organisations**

Both organisations were fighting for organisational credibility by having to consistently persuade other parts of their organisation about the importance and value of good product design. Electrolux regarded conceptual work as providing more challenging and influential opportunities to affect organisational acceptance of industrial design. Whereas in Danfoss, it appeared that greater organisation acceptance of their design methods would only be found once they had been proved on mainstream products.

The MMI group was independently funded and did not have to rely on external design projects for financial support. Electrolux was integrated into the product development process even if this was perceived by the group to provide trivial 'face-lifting' design services. Their ambitions were to move downstream of the product development process and be responsible for high-level strategic design decisions.

User involvement was integrated as a marketing activity in Electrolux to assess marketability rather than product usability. Users helped define market segmentation at the early stages of product development and then were involved at the latter stages by determining market acceptance. The designers at Electrolux were willing to involve users in design activity and provided suggestions similar to those used at MMI.



One of the strengths of the MMI group was the willingness to adapt and improve the effectiveness of the tools used, almost with evangelical fervour. The style and approach of design tools was, however, constantly changing, making it difficult to gain a coherent framework from the interviews. Through the interviews it appeared that defining inputs and outputs from each design tool was difficult. Responses to questions resulted in statements such as 'you get a feel for how to do it'. This made it hard to describe the process as a generalised set of procedural events. Interaction design methods tended to be dependent and reliant on individual or group intuition. Experience and knowledge gained from using the design tools was retained as experiential knowledge. Electrolux, on the other hand, used very conventional design methods that were just beginning to be integrated into their IPD process.

Although Electrolux revealed low user contact in terms of determining design decisions, evidence was found for a willingness to involve users. In contrast, user participation was an essential design driver at MMI and clearly increased the level of dialogue between designers and users. Validity and reliability of these methods was less clear though. Design tools were constantly under revision and group members would adapt them to suit their own needs. Clear measurable objectives or the exact purpose for each design tool was difficult to document. Much of the adoption, adaptation and implementation of the design tools was dependent upon frequent team discussions. Design data were also gathered through experimentation with the design tools thus providing many opportunities to improve the quality of this data. Evidence of generic or formalised design tools was not found. This would make it difficult to transfer these methods to other design groups.

#### **6.4.5 *Potential effectiveness of participative design tools***

While some concern has been raised about the transferability of such design tools, important applicability characteristics were identified that warranted further attention. In order to analyse how effective such methods might be,



the design tools were reviewed using the effectiveness criteria formulated in Chapter 2.

One of the interesting and potentially radical aspects of these design tools was how design data was generated. The *content emphasis* of the design data was grounded within the workshops and captured using a variety of media such as video clips, drawings and mock-ups. Participants and designers produced design data while carrying out interaction scenarios and therefore it was unpredictable and highly dependent on effective communication. This contrasts with passive guidelines, which are derived from external, independently produced design data.

Another important advantage of the MMI tools was the ability to capture context sensitive data; a problem that was identified in Studies 3 and 4. A service engineer, for example, demonstrated how he could design a diagnostic tool. It became apparent to him during a particular scenario that the device must be accessible in very enclosed spaces, which forced him to reconsider the form of the device in more detail. This was captured on video and used as evidence as a 'user requirement'. Video clips were often used to present useful design data. One of the problems of such an approach was how analysis could be carried out using a valid and reliable framework. The members of the design team, for example, did not consistently analyse different abstraction levels of interaction such as the use of a scoping tool such as *ergonomics criteria*. Although some design tools such as 'user characters' and 'mapping' did limit abstraction to user or task descriptions.

The design tools were extremely *relevant to the design of novel interactive devices*. Many of the design tools were developed specifically so participants could consider novel devices and interaction styles such as through role-play or game playing.



Appraising the *usability* of design tools in terms of their effectiveness as a set of generic, transferable design tools, was less apparent. Formal recording or documentation of the design tools was not undertaken and design tools were improved or revised on an ad hoc basis. The design tools were designed by the design team and therefore naturally mapped onto their own working methods. Designers were both producers and users of the design tools. High importance had been placed on producing design tools with strong participant involvement so naturally they had to be engaging. It was obvious from the study the participants using the design tools did find them enjoyable. The core principles, structural framework and theoretical underpinning of the tools remained implicit. For these design tools to be usable elsewhere these issues needed to be addressed.

Specificity Criteria	Applicability Criteria
Content emphasis - Entirely 'empathy driven' by participants and designers together	Usability of design tools - difficult to measure as the tools were designed and used by the designers
Structure and presentation - Wide use of media, high volume of data makes it difficult to select, analyse or capture consistently at same level of abstraction	Reliability and validity of design tools - concern was raised due to the tools being intuitively adapted and evolved
Relevance of design tools to 'ergonomics criteria' - not considered	Relevance to proposed target audience - high because they were developed by themselves
Relevance of design tools to novel interactive devices - specifically designed for them	Likelihood of survival or usage in organisational or design context - evidence of scepticism in other parts of the organisation

Table 6.7 Comparing MMI design tools against effectiveness criteria

The *reliability and validity* of such methods has been highlighted. Handling and controlling outputs (user comments and observations for example) from the design tools appeared to lack any systematic rigour. This problem was raised and it was argued, quite strongly, that findings must emerge through 'participation empathy'. Any form of procedural or analytical framework would inhibit this process. One of the reasons given for having design



workshops in quick succession was to retain intangible and implicit concepts held by the designers and participants from one workshop to another. This was regarded as an important characteristic of the design tools resulting in a clear resistance to formalisation. Subtle, non-articulated requirements could be captured without any formal recording of them.

The *likelihood of survival* of these design tools seemed assured as long as independent funding for the group remained. For these design tools to survive in other design organisations, such as Electrolux, clear benefits would need to be obvious from an early stage. The relevance of such design tools to their *target audience* appeared to be strong. They supported experiential learning and appear to map with important design behaviour found in Studies 2 and 6. They also integrate well with existing design methods such as sketching and mock-up building and require practical and team-working skills. A summary of effectiveness of these design tools is presented in Table 6.7

## 6.5 Conclusions

Study 7 suggested user-centred design was still a minority practice but users were consulted in quite a high number of consultancies. This was identified as a potential area for developing an alternative bridging mechanism. Study 8 suggested that other bridging mechanisms could be developed that were closer to the working methods of designers and involved end users. The MMI design tools demonstrated strong effectiveness in terms of applicability by mapping their design tools to design behaviour. The MMI design tools did not rely on a corpus of external, passive design data but created their own body of data every time the tools were used. However, some concerns were raised about the transferability of these design tools into other design organisations, such as Electrolux for example. Study 8 raised some important questions. As the MMI design tools were used as *primary* tools (indigenous design tools), how transferable would they be in other design groups as *secondary* (externally produced) tools? Is it possible to extract key



principles and theories underpinning these design tools to form a rationale for secondary tools? Can any adapted or refined secondary design tools address the effectiveness deficiencies already raised?

A common theoretical model was required that addressed issues related to controlling design data abstraction and the reliability and validity of gaining data in this way. To achieve this, a second literature review was conducted in the following chapter to build a theoretical model of practice, applicable to interaction design methods that may suit organisations such as Electrolux.



# 7 Exploring the theory and application of context in design

---

## 7.1 Introduction

Accumulating evidence from Studies 2, 5 and 6 concluded that guidelines are rarely appropriate for the intended practice domain of interaction design. Studies 3 and 4 added to this evidence by concluding that proactively creating design data becomes more difficult as design 'novelty' increases. Alternative bridging mechanisms needed to be sought and findings from Study 8 suggested that user-participative and context dependent design tools may overcome many of the applicability and specificity problems identified in the previous studies.

This chapter begins with brief description of a range of theoretical models and applied analytical methods and techniques concerned with context, design and activity. The prototypes developed in Study 3 were used to exemplify these theories and to help illustrate their significance to identifying contextually based interaction design problems. These theories and applied methods are then critically reviewed in order to produce design requirements for the proposed secondary design tools. This process helped to re-define the research context in which the proposed design tools were developed.

## 7.2 Aims and objectives

- To compare theoretical models and existing methods and techniques considering the use of context in human behaviour and design activity to assist in developing design tools requirements.
- To re-evaluate the research framework produced in Chapter 1 (p 10)



### 7.3 Theoretical and applied approaches to the study of context and design

In this section, the theoretical and applied approaches are briefly described and then critically compared in Section 7.4.

It was felt important that brief descriptions were provided as the theories and applied methods formed the basis on which the proposed design tools are based. These theories and applied methods were selected because many of them have had a recent profound effect on introducing new methods of scientific inquiry into HCI (Monk *et al* 1993).

Four theoretical approaches to the understanding of human behaviour within a 'context' (defined as past insights, common intuitions, shared understanding) are presented. These included: Activity Theory, Distributed Cognition, Situated Activity and Reflection-in-action. The fourth theoretical model addresses the use of practice-based knowledge and skills in practice-based activity and is often related to design activity. From this standpoint, methods and techniques that apply these theoretical principals are discussed.

The study of 'context' was addressed from two points of view. First, context can be viewed as part of the design activity. Here context is considered as being part of the phenomenon in which a designer attempts to solve a design problem. Context was assessed in this way in Studies 2, 5 and 6. Second, it can be viewed as a component element of interaction between a user and a product and was addressed in this way in Studies 3 and 4. It is important that this distinction is made between these two points of view of context and design as these are both discussed in this chapter.

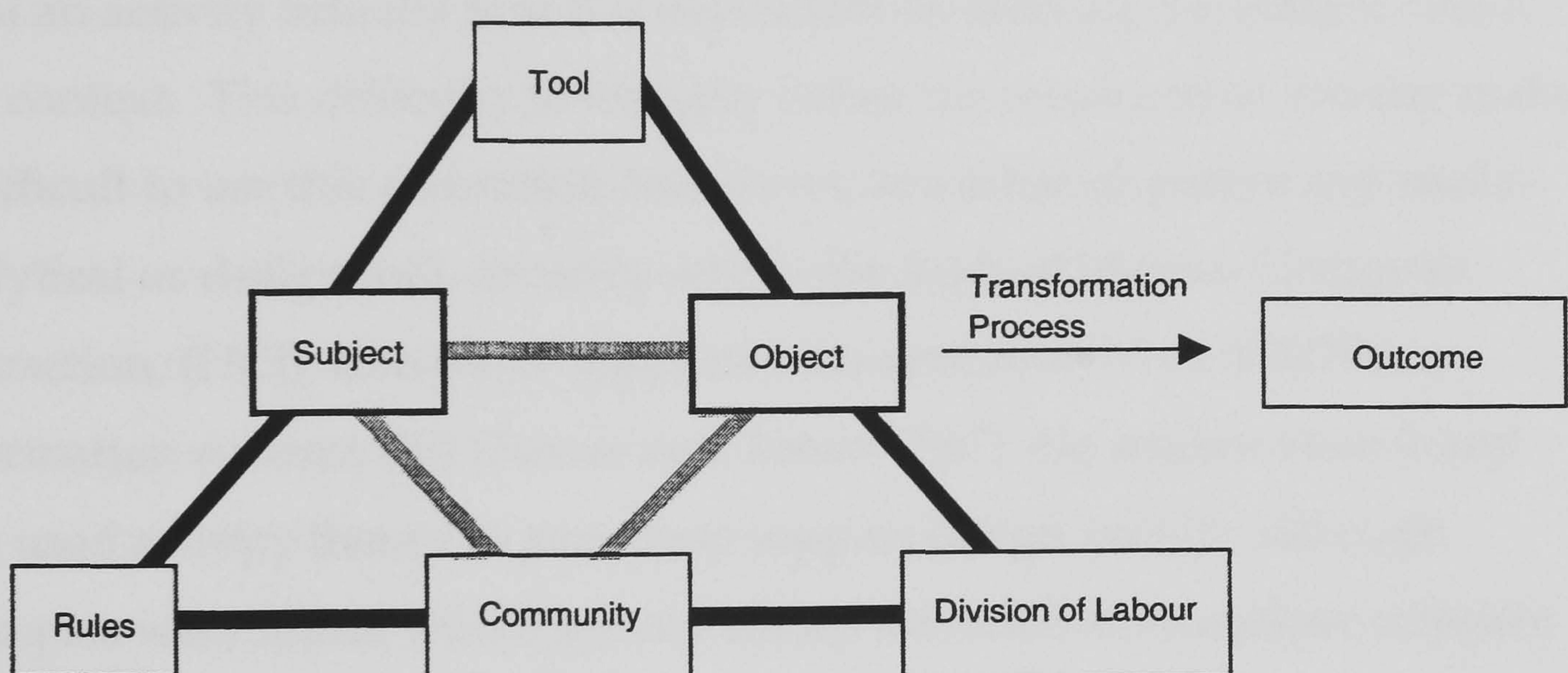
#### 7.3.1 Activity theory

Activity theory was developed by the Russian psychologist Vygotsky (1978) and continued by Leontjev (1981) and introduced to the Western world by Bødker (1991). One of the key principles of activity theory is to bring



together activity, using a wider social and organisational context combining this with 'consciousness' which can be defined as unifying attention, intention, memory, reasoning and speech located within everyday practice (Vygotsky 1978 in Nardi 1996). An activity involves a subject, which can be a person or group interacting with an 'object' (see Figure 7.1).

The term 'object' is an important concept and has a different meaning to its typical understanding in HCI or design. An object can be physical, such as a book, or it can be an idea or concept explained within a book or it can be a shared vision between individuals and can change during the process of activity.



**Figure 7.1 Basic activity structure (Kuutti 1996 p28)**

A subject (the individual or group under observation) uses 'tools' as mediating devices in order to accomplish transformation. Tools can be enabling by improving the effectiveness of obtaining an outcome but they can also, by their form and function, restrict access to the object. To capture context more broadly, relationships between the subject and a 'community' (who share the same object) are considered. This creates two new



relationships: subject-community and community-object. Both these relationships are also mediated.

An activity is therefore a minimal meaningful context that can be encompassed within this structural framework. An activity is not necessarily accomplished as one uninterrupted process but typically consists of steps or phases. Intentional activity therefore has a hierarchical structure with subordinate behaviours known as actions and operations.

Activity theory provides a descriptive framework in which to analyse purposeful, contextually based behaviour. This framework permits flexibility in terms of describing activity but also makes it difficult to define what an activity actually is as it is dependent on defining the subject, object and context. This difficulty to formally define the constructs of activity make it difficult to use this theoretical framework as an intrinsic part of any useful analytical or design tool, certainly within the fields of Human Computer Interaction, (HCI), Computer Supported Co-operative Work (CSCW) or Information systems (IS) (Turner and Turner 1997). No studies were found that used activity theory to *proactively* support design activity although examples were found where activity theory was used to re-analyse software design activity (Nardi 1996).

### *Applied example*

Using the 'drag and drop' interface, Figure 7.2, as an example (for description of interaction style see section 3.4.3), the *activity* can be described as programming the washing machine.



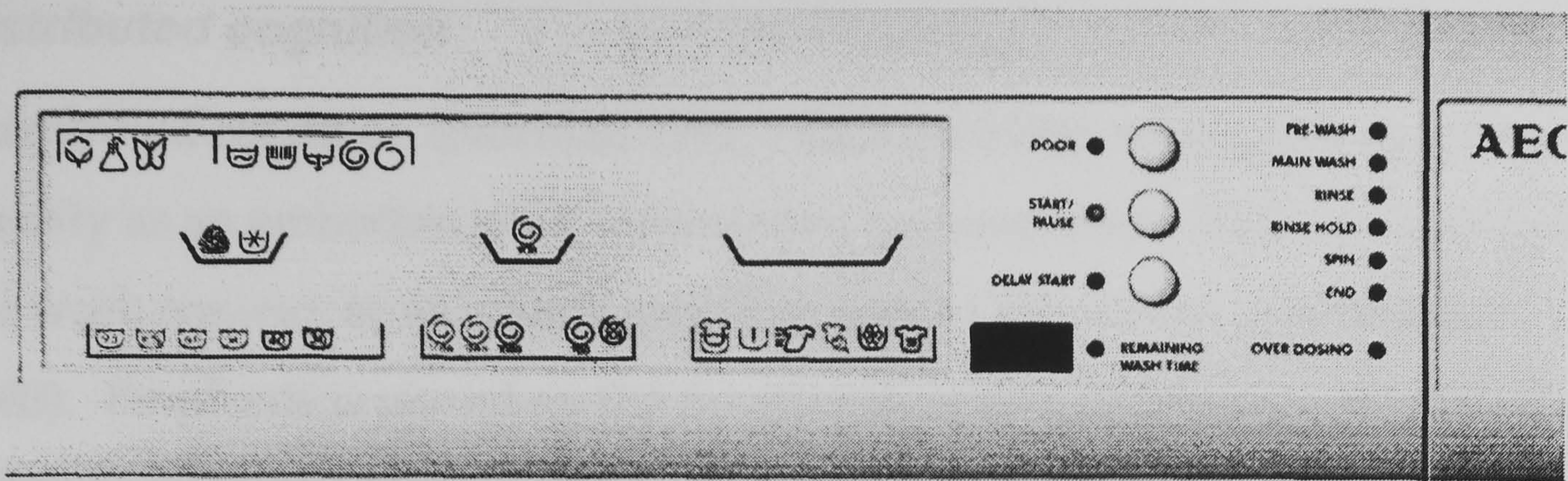


Figure 7.2 Drag and Drop interaction style example

The *object* can be assumed to be configuring the washing to carry out the correct functions. The *tool* used to achieve this *outcome* is the drag and drop interface. As this is a novel interface there are few *rules* the subject can rely upon through *community* relationships. Shared or common experiences from observing other subjects are not available in order to suggest how to carry out correct interaction behaviour. This requires the user to depend more heavily on *operational* level support, visual cues that suggest how the device might be operated. Once the *subject* has learnt that icons can be manipulated and dragged to washing baskets, they might observe machine automated *operations*, the default 'spin speed' jumping into the second washing basket once a fabric type and temperature has been selected. The *tool* is enabling completion of the *transformation process* by flashing a red light next to the start button. As the *subject* becomes more skilled, the *tool* supports further *actions*. The *subject* may quickly recognise a washing program through spatial selection of icons rather than through pattern recognition.

In this example, using activity theory to analyse the 'drag and drop' interface suggests that more visual cues will be required to initially understand the syntax. This was found to be the case in Study 3. Analysis also suggests spatial cues might offer good action support with prolonged use.

Unfortunately this could not be tested in the evaluation study due to the short duration of the experiment.



### 7.3.2 Distributed cognition

Distributed cognition (Norman 1991, Hutchins 1994) emphasises cognitive activity as an embodiment of information between users, their artefacts and the work context, or as a 'joint cognitive system' (Hollnagel and Woods, 1983). Emphasis is placed on the functional aspects of artefacts and how they relate or affect the task (as opposed to the more broadly defined term activity used in activity theory). Artefacts can change the characteristics of a task by distributing actions across time and individuals. A checklist, for example, can be prepared and written once and then subsequently used many times. A main goal of analysing distributed cognition is to observe how information is propagated throughout the 'functional system' by using cognitive devices. The unit of analysis is a cognitive or functional system composed of individuals and artefacts in use. Analysis is focussed on how information moves and transforms between different cognitive artefacts and how well artefacts support actions towards a goal, also how well the system interprets differences between intended and actual action.

#### *Applied example*

The 'animated object' interaction style (Figure 7.3) demonstrated how the surface representation of three combined functions, the animated object, simply did not adequately transform the underlying representation of three functions (spin speed, temperature and fabric type).

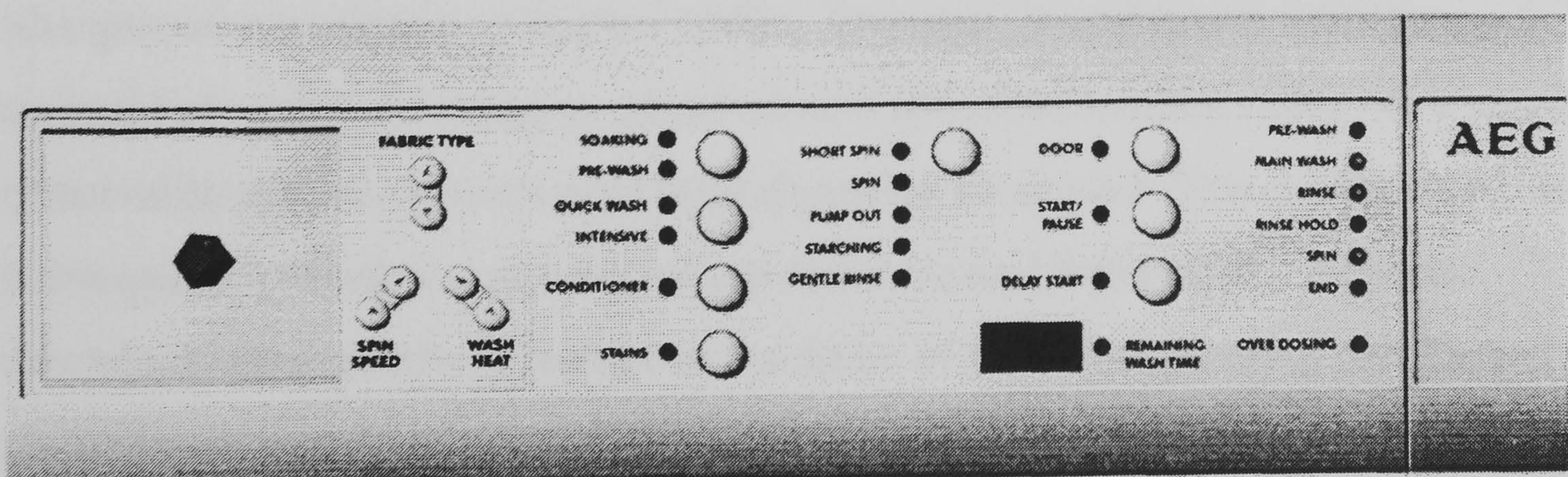


Figure 7.3 Animated Object interaction style example



The animated object was designed to quickly 'suggest' how the three functions were configured. Representation between surface and internal representation was too tenuous and simply did not support the intended goal. The surface representation was too abstract and lacking in contextual relevance to convey the combined relationships between the three functions.

### **7.3.3 *Situated action model***

Much of the work on situated action was brought to the attention of the HCI community by Suchman (1987) where she argued that actions can no longer be regarded as the simple execution of pre-established plans. In cognitive science, it had been argued that the organisation and significance of human action is based on underlying, predetermined plans, where plans precede action and, if conditions change, new plans are engaged.

Suchman challenged these assumptions about purposeful action and shared understanding by using methods from the social sciences. In situated action, the relationship between knowledge and actions are part of the analytical study of purposeful action. Observational studies establish how people develop and derive plans during situated action. Actions depend upon artefacts and social circumstances and are not subordinate to plans and can not be abstracted from the situation under observation. Plans may be engaged before action takes place but during the action, plans are often abandoned resorting to 'situated' skills. The unit of analysis is activity embedded within a setting or situation and involves inquiry using very small observable actions (glances and hesitations for example). This is necessary, as subsequent interviews may not reveal true situated behaviour. Actions described retrospectively maybe in the form of intended plans or recollected reconstructions.



During action, an event known as a 'breakdown' may occur and these are important for analysis. Breakdowns occur when the tool-action relationship becomes disturbed in some way, through some unexpected or unanticipated outcome such as an error message on a display screen. This forces new plans to be initiated. In situated action, it is argued that these rules and procedures are not self-contained previously developed plans but plans for the purposes of deliberation to re-evaluate action. For this reason studying situation action relies on considering minute actions and utterances which can often be ignored when presented through a descriptive or observational account. Analysis is aimed at defining the relationship between interpretations of actions and the action's circumstances.

### *Applied example*

Situated action provides the possibility of a strong analytical approach. Although analysis was not undertaken at this level of detail in Study 3, its approach can be exemplified through the auditory display interface. Teachback protocols demonstrated how plans become situated. Participants could not easily distinguish between variation in pitch to select spin speeds; many participants adopted a plan of repeatedly hitting the 'spin speed' button until they reached the highest or lowest pitched sound and then counted increments up or down to obtain the correct spin speed selection. This was an obvious and explicit strategy and because of this it was captured in the teachback protocols. However, there was evidence of other situated plans not captured in the teachback protocols (omitting repetitive elements of the task for example) to illustrate that situated activity can only be recorded through detailed observation. This detailed analysis may assist in design activity but requires a robust interaction environment before meaningful findings emerge. Activity is heavily embedded in the situation, if the situation is simulated (as in the case of prototype development) then attempting to describe situated action would be fruitless.



### 7.3.4 *Reflection-in-action*

Schön (1983) presents a contextually sensitive model of professional practice and uses design activity as an example of how a designer uses a mixture of tacit knowledge along with the situated design task to solve the 'problem'. The design problem is framed using the process of 'reflection-in-action'. During this process, the designer engages in a reflective conversation with a unique and uncertain situation, which involves re-framing the 'problem'. The designer conducts experiments to discover consequences and implications and then adapts the situation to the intended frame. This may produce unintended changes, which in turn prompts new meaning. In this way the situation talks back and the designer listens providing new discoveries which then call for new reflection-in-action. Schön argues that to make reflection-in-action a more explicit and formalised part of practitioner (designer's) activity, greater awareness needs to be made of this reflective conversation. Designers need to become aware of their 'tacit frames' and the possibilities of altering the reality of practice. Where new practice is developed from situations that do not fit available routines, known as 'repertoire-building', these should be made more explicit and open to analysis; for example, decision making. Attempts should also be made to explore the mechanisms or triggers that practitioners use to make sense of new situations. Finally, forcing practitioners to examine a situation without the use of conditioned responses can help in analysing reflection-in-action. Schön (1983 p322) states, 'As we try to understand the nature of reflection-in-action and the conditions that encourage or inhibit it, we study a cognitive process greatly influenced by "cognitive emotion" and by the social context of inquiry'.

#### *Applied example*

The QOC analysis (section 3.4.4) provided a very simplified model of reflection-in-action. The designer engaged in a conversation with the design situation by breaking one of the design problems into three stages (behavioural attributes of icons). The inexperience of the designer was



revealed through the narrow range of possible and contingent interaction options that were considered (see Figure 3.6). Much of the research that followed this study attempted to identify how such 'repertoire-building' and 'triggers used to make sense of new situations' should be made explicit. Evidence has already indicated that these factors were not enabled through a passive text-based medium such as guidelines.

### **7.3.5 *Using awareness of context in design***

The theories, briefly described above, provide differing viewpoints on how context influences purposeful activity. Applying these theoretical insights to design, however, can be difficult. This can be illustrated through a special HCI journal issue dedicated to design and context in which Brown and Duguid (1994) offered a descriptive framework to assist in using context in design. In this framework they use the notion of context as having a centre and a periphery. The centre is understood as being the point of concern or attention. The periphery includes aspects of interaction (sounds, supporting artefacts or cultural norms, for example) that do not have a direct influence on interaction activity. However, peripheral issues can move to the centre and vice versa depending on changes in the interaction activity.

Connecting between the centre and the periphery is the notion of a boarder. This has a contingent relationship between the centre and the periphery and different boarder issues may reveal themselves as context is considered. The noise of keyboard, for example, could be regarded as a boarder resource in that it provides important feedback cues to a typist. This resource can also have a 'socially shared' significance. In this example the noise can assist others to know if the typist in an adjacent office is present or not. It is the boarder issues that they suggest designers need to be made more aware of.

However, criticism from invited respondents, who critiqued their position paper in the same issue, were concerned with how designers were supposed to grasp and use this framework and questioned the benefit of a conceptual



model of context. Respondents offered applied methods and three of these are discussed below (contextual design, scenario design and interaction analysis) many of which are based on ethnographic analysis or participative design.

Participative design empowers users by regarding the users as the domain experts and assuming that any changes to a system should improve their role within it (Schuler and Namikoka 1993). The fundamental philosophy of participative design is to involve the users of future systems or artefacts in design activity (Greenbaum and Kyng 1991).

The purpose of ethnography is to obtain a better qualitative understanding of these interactions. Data gathering is through participant observation, where the observer is aware of their role and significance as *part of* as well as *distinct from* a situation under study. Ethnography is a complex and difficult data gathering process where no one method exists (Hughes *et al* 1996).

Ethnographic studies have been criticised for being based on small, confined environments, time consuming and being practically difficult to implement. It has also been found difficult to frame results in an adequate manner for designers (Hughes *et al* 1995).

Rogers and Bellotti (1997) attempted to address this problem and find out what kind of bridge was required between 'blue-sky' design activity and ethnographic data. They examined what type of dialogue would be required between designers and ethnographic researchers to lead to innovative solutions. They state that it was necessary to be able to translate the findings from studies into a form that is accessible by those who have not had the opportunity to go out into the field. A reflective framework was developed using questions that specifically dealt with existing design problems and potential design solutions. By answering these questions by providing data in the form of video clips and photographs and in close dialogue with the designers, prototype solutions were generated.



### 7.3.6 *Contextual design*

Other, more structured, methodologies analysing context and design have evolved. Beyer and Holtzblatt (1998) proposed a methodology known as contextual design specifically tailored toward the development of contextually based design data using much of the theoretical underpinning already described above. Many of the techniques employed in contextual design help to develop a coherent understanding of the work and systems to support work through explicit and sharable design data. The methodology employs a mixture of ethnographic and participative design methods.

The methodology has six stages. At each stage importance is placed on ensuring that contextual information is gathered by making sure that findings are grounded in real situations. Representations of existing work methods and proposed systems are expressed using five different types of work models to ensure that as many aspects of work in context are captured. Focus is placed on improving work methods and not on technological solutions.

### 7.3.7 *Scenario design*

The complexity and subtlety of interaction makes comprehensive descriptions of activity difficult. The use of scenarios as a design tool has evolved in HCI as a mechanism to describe complex activity by allowing users and designers to engage in and articulate design intentions. Scenarios are represented in many different ways such as textual narrative, storyboards, video mock-ups and scripted prototypes. Definitions of a scenario vary but Kuutti (1995) suggests they have two key ingredients as a description of activity: first, as a narrative form and second, as a representation from the viewpoint of the user. The application of scenarios in terms of the software development lifecycle is flexible and varied and can include requirements analysis, user-designer communication, software design, design rationale, abstraction and team building (Carroll 1995).



Examples of scenario design are growing such as using scenarios through co-operative design. Kyng (1995) developed future user requirements involving the use of hypermedia, a concept that the users were initially reluctant to contemplate until introduced to the concept through scenarios. They found that the open-ended aspects of evaluating scenarios allowed features to be considered that would otherwise have been designed out had more formal evaluation tasks been undertaken. Karat (1995) used scenarios to focus discussions and to elicit requirements from potential users of speech recognition technology and also found that not overly constraining exactly what a scenario was helped in addressing broader user requirements.

### 7.3.8 *Interaction analysis*

Video taping provides an opportunity to analyse human activity while considering contextual detail. Jordan and Henderson (1995) describe 'interaction analysis', which has been used in many different research fields, including HCI. The goal of interaction analysis is to reveal how participants use the resources of actors and objects within a particular context and has its roots in such approaches as ethnography. Using video recordings of a situation of interest, a content log is first produced, tapes are then analysed in groups to propose insights and hypotheses. Speculations must be grounded in the empirical evidence. Participants under study can also provide useful data by being involved in the review sessions. Good interaction analysis focuses on the 'object' (in activity theory terms) and not on the subject or the tool.

Lewis *et al* (1996) used an adapted form of interaction analysis to produce design data for a user interface for a video editing software distributed across a network. The researchers quickly found that data capture and analysis methods had to be pragmatic to fit commercial demands. Time constraints and lack of expertise required a more streamlined approach to using interaction analysis. A coding scheme was devised that was efficient with simple nomenclature. A more refined model of observed activity was



developed by evolving the coding of video segments during team discussions. Finally the model was evaluated against its effectiveness at answering previously set requirements criteria and design proposals.

The model then formed a platform on which design proposals were proposed and assessed. They found that the model could not be represented as prose and found the use of scripts more productive where scripts contained a number of scenes, objects and roles. Scripts were then passed to design teams as scenarios as constructed stereotypical stories. Interface design elements were assessed against the task domain it was supporting. In this way the model acted as an interim representation between design solutions and real task activity. It was also important that the model was used by different user groups, such as the UI designers (themselves) and product engineers and be quickly communicable to other groups in the organisation.

## **7.4 Comparing theoretical and applied approaches to design and context**

In this section the theories and applied methodologies are discussed allowing a set of design requirements for the proposed design tools to be put forward. This is achieved by first posing questions on how existing research studies deal with context in design. Common principles and methodological issues from the theoretical models and application of theory are discussed to help in the development of design requirements for effective secondary design tools. After discussing each question, a design requirement is stated.

*Are there any common underlying principles that cut across the theories of activity theory, distributed cognition and situated action that help in constructing contextually sensitive design tools?*

All theories (in Section 7.3) supported the notion of knowledge evolving through action and being social in origin. The organisation and use of knowledge are not limited to internalised mental cognition but in the



interaction between actors, artefacts and environments. Interaction with artefacts and environments creates contextually embedded knowledge. But how does an understanding of this paradigm improve the usability of CPT? This paradigm is certainly having an impact on HCI research, as evidenced by the literature review in this chapter, but the recognition or importance of these social science-based theories equates more to a re-defining of interaction analysis. This does not necessarily improve or make the process of designing interaction more obvious or help identify usability problems. The review was less clear about how these theories can be directly applied as analytical tools to assist in interaction design.

Using the theoretical models to identify contextually based usability problems was possible by re-evaluating the prototypes used in Studies 3 and 4, although the effectiveness of this must be taken with caution as it was done retrospectively.

Context can only be realistically analysed through detailed ethnographic based studies but this can not be assumed to be viable as designers are not trained or able to conduct such detailed and intensive ethnographic studies (see Studies 6 and 7). Other methods of collecting design data were observed which appear to offer useful bridging mechanisms. Studies by Beyer and Holtzblatt (1998); Lewis *et al* (1996); and Rogers and Bellotti (1997) suggest that contextually based design data can be gained in an attenuated form where detailed ethnographic studies are adapted to focussed structured analytical methods. Furthermore, observations in Study 8 tend to support this argument.

### **Design requirement 1**

**Context based design data should be gathered but will have to be in an attenuated form.**



*Do applied and analytical methods found in the literature suggest how context can effectively be captured as an integral part of design data?*

Jordan and Henderson (1995) offered many suggestions on data analysis techniques but the process of producing contextually rich data was didactic and time consuming. Interaction analysis offered a procedural framework but did not offer an obvious route to developing design data from complex grounded data, although this is understandable as it was developed as a generic analytic method to be across different disciplines.

Many of the applied methodologies overcame this problem of generating design data by constraining the contextual problem and converting findings into a representation model. Beyer and Holtzblatt (1998) used five different abstraction models to handle the complexity of capturing context. Both Lewis *et al* (1996) and Beyer and Holtzblatt used a form of intermediary model first to evolve an appropriate and robust abstraction of the situation under study but also to test design proposals against it. Whereas Rogers and Bellotti (1997) use a design-based inquiry approach to gather data that designers would find beneficial. These studies suggested that using a transitional approach from moving from grounded data to design data was possible if some form of representational model of interaction activity was used. Proposed design tools must both support a representational model of interaction that supports design activity and should act as an intermediary between design data and interface design proposals. Any proposed model or procedures to develop such a model must articulate clearly what the unit of analysis is to be and should indicate what should be recorded (Suchman 1987).

## **Design requirement 2**

### **Design tools must create and support representational models**



*What type of representational model of interaction should be used?*

Contextual design and scenario design use many forms of representational models such as rough sketches and 'post-it' note diagrams, video mock-ups and storyboards. Users and designers should naturally understand good representational models. Ehn (1992) suggested the use of common 'language-games' between users and designers. These do not necessarily need to make the same sense to users and designers, only that the rules of participation make sense. Design artefacts should not attempt to create 'pictures of reality' but should help users and designers articulate current situations and envision future ones. This should be achieved through 'design-by-doing' using mock-ups and prototyping artefacts where behaviour and shared understanding is apparent to both designers and users. This is a natural approach that designers adopt. Study 2, 7 and 8 provided evidence that designers prefer to employ mini experiments rather than seek out corroborating design data. Designers also find it difficult to frame problems outside of their experiential or tacit knowledge (Study 5) and tend to learn from reflection-in-action. Tools should also provide mechanisms to allow designers to reflect on their own learning (Schön 1983, Ehn, 1992).

### **Design requirement 3**

**Interaction models must offer a common language (e.g. design-by-doing) which is quickly understood by designers and potential users**

*What unit of analysis should be used for contextual analysis?*

Activity theory and situated action have units of analysis at very different levels of granularity. Activity can be concerned with broad social and cultural issues whereas situated action is more embedded in a specific situation. Both would appear, on the findings from Studies 6 and 7, to offer analysis at too high or low a level, as contact with users is cursory and limited. Cognitive distribution, on the other hand, uses artefacts within a



cognitive system, which intuitively appeared to map onto the type of problem domain that interaction designers are likely to face.

In scenario design, defining boundaries of interaction is often vague. The problem domain is contained within a scenario and it is argued this vagueness of abstraction definition assists in creatively analysing the situation-at-hand. Selection and granularity of scenarios, however remains a problem. Solutions are not offered about how effective scenarios should be selected or developed. Mack (1995) argues that scenarios can only be improved by iteration using interactions between system users and the designers.

#### **Design requirement 4**

**A suitable unit of analysis should be obtained through subsequent evaluation studies of the proposed design tools**

*Can the context of design activity be exploited in the proposed design tools?*

Reflection-in-action provided a useful insight into the use of 'conversation' between the situation and the problem frame and Schön offered suggestions on how these dialogues can be made more explicit. Schön identified factors such as repertoire building and the use of triggers as mechanisms to bring reflection-in-action to the fore. Evidence from Study 8 revealed that designers were happy to evolve and adapt their design methods to improve their overall effectiveness and, in this way, they were openly exploring the conversation between the situation and the problem frame. Although, the extent to which their design methods evolved in a structured or formalised way was unclear from the study.

#### **Design requirement 5**

**Adaptation of design methods should be considered as an integral characteristic of any proposed design tools.**

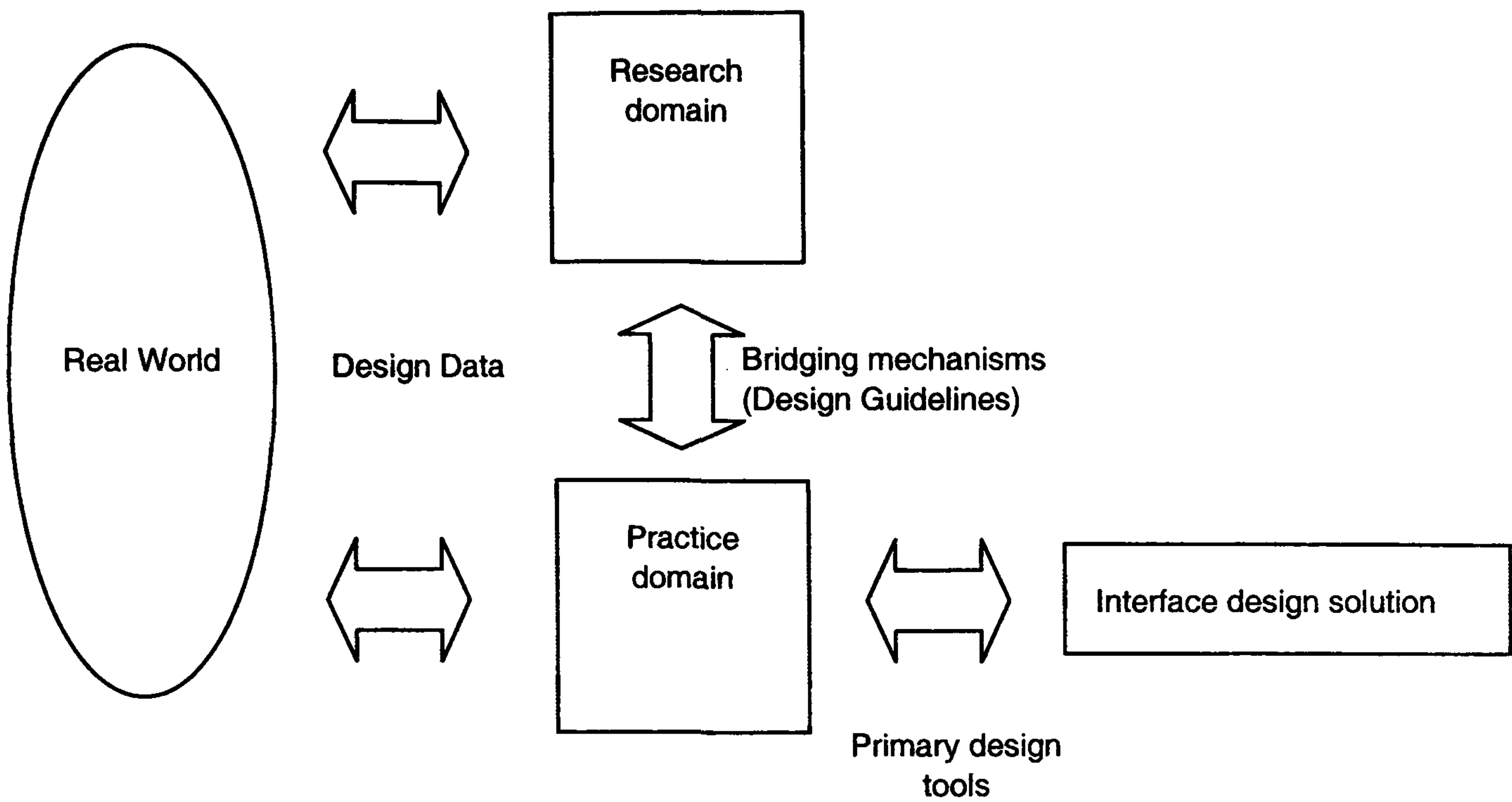


Answering these questions helped in beginning to formulate the design requirements for the proposed design tools embracing contextual analysis and participative design methods. Reviewing the theoretical foundations and applied research based on these foundations helped to provide direction. The most important constraint that had to be imposed on the design of future design tools was they needed to be used as *secondary* design methods. Evidence from studies 2, 6 and 7 suggested designers were not prepared to adopt or apply new design methodologies readily.

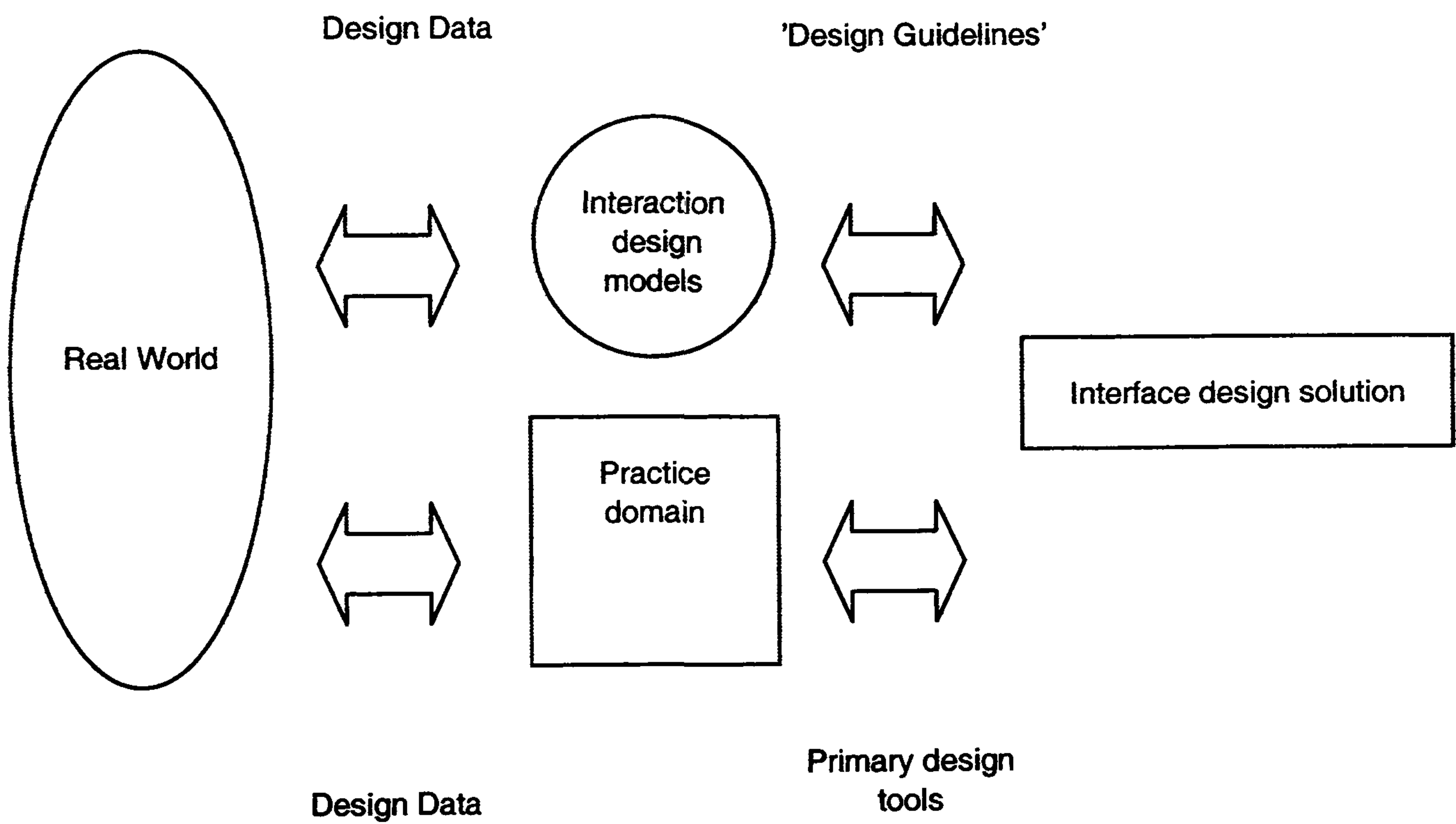
## 7.5 Design tool framework

These design requirements were used to re-visit the research framework presented in Chapter 1 (Figure 7.4). To remove the need for any form of external, non-bespoke and passive bridging mechanism, it was proposed that context sensitive design data should be generated *in parallel* to conventional design activity through the development of a series of *interaction models*. Interaction design models would act as a transitional process between gaining design data from the real world and producing problem specific design guidelines which could then support usable interface design solutions (see Figure 7.4 b). The term 'design guidelines' was retained but it was not anticipated that they would resemble guidelines used in the previous studies but would be in the form of tangible design guidance. The intention was that the proposed secondary design tools reflect, as far as possible, primary design tools used in the practice domain. Three design tools were proposed. The initial design tools are described in detail in Chapter 8 along with an account of how they evolved through iterative development.





(a) Original Research Framework



(b) Revised Research Framework

Figure 7.4 New framework to gain design data and guidelines



## 7.6 Conclusions

In this chapter a narrowly focussed literature review was carried out to address the issue of context and design. Descriptive and analytical models of context were reviewed which helped to define and propose secondary design tools that incorporated contextually rich design data. Five design requirements, based on findings from the literature review and the previous studies, were produced and were used to redefine an alternative framework in which the proposed secondary design tools could be integrated.



## 8 Iterative development and evaluation of interaction design tools

---

### 8.1 Introduction

Theoretical and empirical evidence provided a strong and compelling argument that design tools could assist interaction designers and potential users to collaboratively conceive, develop and evaluate novel interfaces for Consumer Product Technology (CPT). However, the methods and techniques reviewed in the literature in Chapter 7 were carried out by researchers with extensive knowledge of the theoretical background and had appropriate investigative skills. These methods and techniques were being used as *primary* investigative tools. The challenge in this chapter was to investigate if such methods and techniques could be adapted to form a set of *secondary* design tools. The important, and unknown step, was to establish if such methods could be used and adopted effectively by designers who were not trained in ethnographic and user participative based methods.

Action research methods were used to adapt, evaluate and improve the design tools. It was intended that iterative studies would be carried out at Electrolux but unfortunately this was not possible due to commercial project demands. Therefore, Studies 10 and 11 were carried out in more controlled, but less contextually sensitive, environments.

### 8.2 Aims and objectives

- To establish whether the initially proposed tools could be used as secondary design methods
- To iterate development of the proposed design tools to form a coherent and inter-related set of stand-alone interaction design tools specifically tailored to novel CPT interfaces



- To establish a revised set of effectiveness criteria and assess the design tools against these

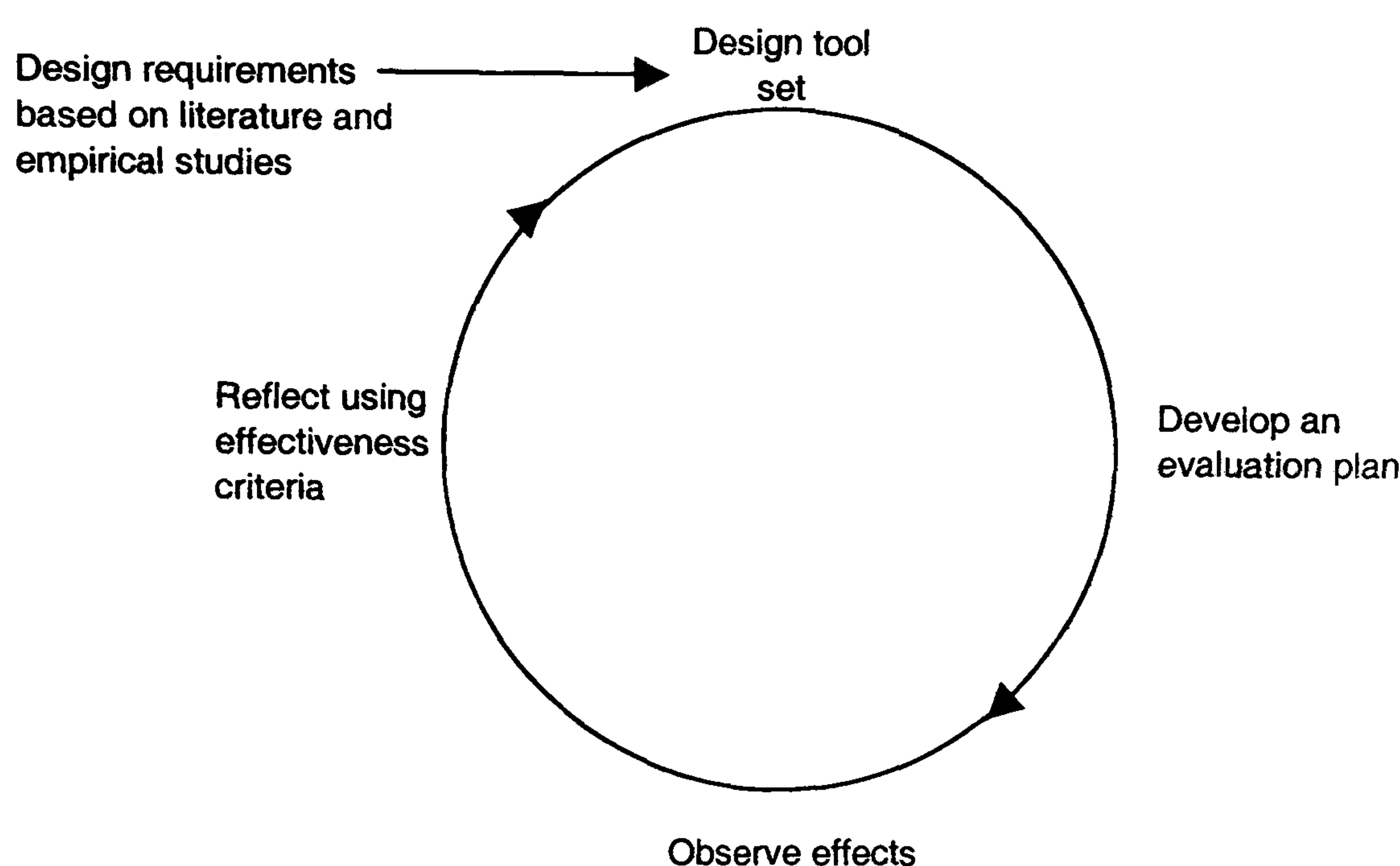
### 8.3 Methodology

Introducing and evaluating new design tools presented a host of methodological problems and uncertainties. Many of the detailed procedural elements of the proposed design tools had yet to be defined. Broad design requirements could be offered, but there still remained a level of uncertainty about whether these requirements would be suitable for secondary rather than primary design tools. To carry out these investigations, therefore, an action research approach was adopted. This allowed investigation to be carried out through participative and observational involvement with the potential users (interaction designers) of the proposed design tools. Implementing and evaluating the proposed design tools during a live commercial design project was also a concern. Apart from the logistical difficulties of introducing the design tools, care about the design of the participative and observational study needed to be taken. Differentiation between two types of intervention needed to be made: designing the 'design tools' themselves and designing and obtaining solutions from the 'design problem'. For example, altering the level of design tool ownership between the researcher, designers and the user participants may affect how the design problem (interface design project) was framed or analysed. Conversely, the design problem may restrict or affect how the design tools could be implemented. Until the design tools had actually been implemented, it was difficult to know what balance needed to be achieved between intervening in designing the tools and the problem. It must be stressed that, while the design outcomes from the design tools was important, the key emphasis in this study was to improve the effectiveness of the design tools using the effectiveness criteria.

An action research approach was selected, as these types of 'interventional' problems are common in other fields of study, particularly education (Zuber-



Skerrit, 1992). Action research provided an approach where complex and inter-related problems could be observed and reflected upon using qualitative data gathering methods. Issues were allowed to emerge from these observational data as it was collected.



**Figure 8.1 Action research model for design tools evaluation**

A set of design tools, based on the design tool requirements, was developed and is described in Section 8.4. An evaluation plan was then produced where the proposed design tools could be observed during a live commercial design project. The effect and effectiveness of the design tools was observed and assessed resulting in a revised set of effectiveness criteria. Improvements were then made to the design tools in light of reflections based on the revised effectiveness criteria (see Figure 8.1).

## 8.4 Rationale for design tools

A detailed rationale for the first iteration of design tools is provided, based on the design tool requirements derived in Section 7.4, findings from Study 8 and related tools and techniques found in the literature. Three design tools were proposed and are described in detail below. In supporting a rationale for these design tools, numerical references are made to the design tool requirements (DR) described in Section 7.4.



### 8.4.1 Card sorting

The intention of the first design tool was to allow an interaction design problem to be understood by both designers and prospective users. The intention was that the design tools would gather broad, speculative factors such as roles, objects and procedures within a contextually defined scenario using a proposed CPT interface (DR1). Participants and designers needed to gain a shared understanding of a design problem (DR2). The intended outcomes from this tool should be a range of high-level user requirements to help define a range of design solutions. The tool had to be quick and easy to learn and implement (DR3) and adaptable (DR5).

Observations from Study 8 and further evidence from the literature suggested that 'card sorting' games might be a useful technique to satisfy these requirements. Card sorting games have been used by Muller *et al* (1995), the first stage known as CARD (Collaborative Analysis of Requirements and Design) uses cards to facilitate the articulation of task and communication within working groups. The CARD approach has been modified by Lafreniere (1996) in the design of computer based user interfaces (CUTA) to enable a simple, user derived, task analysis to assist in interface design. Both the CARD and CUTA methods were based on cards depicting elements of tasks activity such as tasks objects, for instance telephones and notepads, process based activities like methods of working and situations; participants within the task activity were also depicted. Both methods require the participants to select task elements and place the cards in an agreed plan or sequence. Once complete, an agreed summary of their 'representation' is given by the participants.

The attraction to adopting this form of technique was the ease with which such a method could be learnt and implemented. Other methods such as role playing were considered but these appeared to require more collective team commitment to be successful. It seemed that a card sorting exercise might be more acceptable and controllable at this early introductory stage, where



designers might be sceptical about using such methods. A series of small incremental card sorting exercises were proposed for the first tool and would be carried out in a workshop setting with designers and potential end users (participants). These exercises were adapted from the CARD and CUTA methods. It was hoped the designers would quickly grasp the procedural aspects of each sorting exercise to build a model of interaction. Each exercise was designed to capture specific contextual design data.

Cards would be designed prior to the workshop by designers to depict contextually important events, tasks or task elements, objects or artefacts and individuals. The descriptors on the cards needed to be carefully composed by the design team to ensure appropriate discussions took place. Depending on the objective of the exercise, the intention was that participants would arrange the cards, adding new ones if necessary, which fitted their jointly agreed representation of the scenario. It was hoped that the arranging of cards would prompt discussion between the participants both about the card content and also where they should be placed on the table. In this way, a scenario or activity could be represented using card placement and content along with discussions and notes taken during each exercise. These outputs would therefore form design data. It was hoped that the cards would allow users and designers to think creatively about the scenario, permitting an open discussion about their needs for a proposed CPT interface.

### *Design of cards and exercises*

Four different types of card sorting tasks were designed each having a different contextual focus. As little was known about the power and impact of this design tool on participants and designers, a 'bottom-up' approach was adopted with the first exercise drawing on concrete task activity before embarking on more abstract concepts.

The first exercise required the participants to prepare and cook a meal using a series of cards describing cooking activities such as 'Check carrots to see if



they are ready'. Other supporting activities were described on cards, such as 'Check to see if oven temperature is correct' using minimal references to technological support as possible. The intention was to allow the participants to discuss the whole cooking process and allow their personal habits and attitudes towards cooking to emerge. Participants lay cards out on the table in the form of a *task plan* describing how they would cook this particular meal.

The second exercise was designed to build on the first by inserting *function cards* containing descriptions of cooker features or technological support. These cards were divided into three groups: low, medium and high level features. These groupings were colour coded. Low level features included heater controls and selector switches, while medium level features included auto timers and temperature probes. High level features included cooking and menu planners. Participants would be instructed to add features to their task plan only if they felt it would be used. Participants would be asked to openly discuss the advantages and disadvantages of each feature before adding a 'function card' from the task plan or deciding to leave it.

The third exercise enabled participants to think about a week in the life of a cooker and place cards depicting other cooking activities, 'clean the cooker' or 'cook a quick snack' under cards labelled with the days of the week. The purpose of this exercise was to explore if participants would be able to make design decisions or make inferences from their cooking habits that might affect more broader or non-task specific interface design issues. The intention of this exercise was to build up a frequency profile of usage and the type of tasks undertaken during a typical week.

In the forth exercise, participants were provided with a series of cards containing character profiles describing fictional individuals with different levels of interest in cooking and technology. Participants would 'match' these characters against some of the function cards used in exercise 2. The



intention for this task was to examine if users could make 'third party' design decisions on behalf of fictional characters.

The card sorting tool provided four main advantages. First, cards would provide a quick and cheap discussion mechanism or act as 'transitional objects' allowing more critical contextual thinking to occur about product interaction. Second, by providing a broad range of card descriptors they could allow novel concepts to be introduced without having to design the interface to support such a concept, allowing participants to interpret or define the cards on their own terms. Third, card descriptions could be divorced from defined or existing technology so future functionality could be discussed. Finally, the placed cards could act as a conceptual 'interaction model' for analysis.

There was no need at this stage to have a single or coherent solution. The intention of the design tool was to arrive at a series of creative concepts that could be used to consider new forms of interactive technology. The development work for this design tool was carried out in Studies 9 and 10

### **8.4.2 Inspection based evaluation tool**

One of the initial concerns of the card sorting tool was its effectiveness to abstract all aspects of an interaction design problem. There was the possibility of building 'error' into this type of method. Woods *et al* (1996) addressed the question of why user-centred design products fail the intended target users and identified a gap between designers' intentions to be 'user-centred' in approach and actual practice. They suggested that organisational constraints force design methods to meet tight deadlines, demonstrate progress and use limited resources. These constraints restrict design teams, preventing the examination of design problems using 'second order iterations', which are learning processes that allow iterations within iterations.



Cognitive dimension	Description
Abstraction Gradient	Does the number of required abstractions create a slight or uphill struggle for the user?
Closeness of mapping	How close are the system's concepts for objects and actions to the user's (likely/assumed) concepts?
Consistency	Can knowledge of other parts of the system be exploited to guess/understand features new to a user?
Diffuseness/Terseness	How (in) efficient (diffuse/terse) are typical/key interactions?
Error-proneness	Can users make errors which good design would prevent?
Hard Operations	When must users work hard during interaction?
Hidden dependencies	Can a user see where displayed values come from?
Premature commitment	Must user provide inputs or make decisions earlier than the task would demand?
Progressive evaluation	Can users monitor towards successful task completion?
Role-expressiveness	Can users determine the purpose of display elements and how they relate to the system as a whole?
Secondary notation	Can users provide input that will not be interpreted by the system?
Viscosity	How much work must the users do to effect change?
Visibility	What must users do to find information?

Table 8.1 Brief description of cognitive dimensions

This design tool was therefore proposed as a form of second order iteration to formally highlight aspects of interaction that have not been considered or abstracted from the card sorting tool. Any aspects that needed to be considered could then be addressed within another iteration of card sorting exercises.

Inspection or expert based methods have received wide acceptance within the field of HCI. Inspection methods rely on a human factors specialist working through an interface using an analytical framework while assuming to be a 'representative' user (DR 1 and 2). Methods such as heuristic



evaluation (Nielsen 1994) and cognitive walkthrough (Polson *et al*, 1992, Wharton *et al*, 1994) have been developed to assess the usability of an interface design proposal when it is still being designed. However, these methods assume that the interface under evaluation is complete and uses traditional interaction styles such as GUI.

This tool needed to be used by non-usability experts (DR3) while addressing novel interaction styles. This placed challenging demands on any form of inspection based method, as it would have to be used beyond conventional intentions. One inspection based method offered some degree of flexibility and might overcome these demands. Cognitive dimensions (Green and Petre 1996) offered a 'broad-brush discussion tool' to help designers, users and prospective purchasers or customers to talk about a system. These cognitive dimensions (CD) serve a similar purpose to 'ergonomics criteria' (see Section 2.5.3) but appeared to be more forgiving in addressing novel or innovative interaction styles and had been developed using 'checklist' criteria whereas the 'ergonomics criteria' had not. Designers evaluate an interface using 13 'dimensions' (see Table 8.1), for example 'viscosity' (allowing designers to think about the ease with which a user might move from one system-state to another).

Once a range of outline design proposals had been developed from the card sorting tool, the intention was that outline proposals would be assessed using the cognitive dimensions. A checklist would be provided to designers based on the dimensions and the intention was that usability problems not considered or encountered during the card sorting would be identified. Considerable work was required to develop a checklist format that would suit non-usability experts and novel interaction styles. This work was carried out in Study 11.



### 8.4.3 *Scenario design tool*

The rationale for this design tool was to allow designers and users to refine and synthesise the design proposals, developed through card sorting and evaluated through some form of CD checklist, using a rich contextual evaluation scenario (DR1). A scenario based design method appeared the most attractive option due to its flexibility and openness to interpretation. Other methods such as contextual design and interaction analysis were too structured and would require high learning costs to the designers. The scenario design tool would permit designers and participants to explore proposed concepts while acting or role playing within a selected scenario (DR2 and 3). The intention was that participants would be able to make more informed and context sensitive judgements about the range of design proposals suggested in card sorting exercises. In order to make the design tools adaptable (DR5), participants were to use paper-based prototypes to enact their activity and walkthrough the interaction procedure. Amendments would be made in discussion with designers where possible new prototype variants could be introduced into the scenario rapidly.

Little practical evidence had been found on how best to implement such a design method and was therefore the least operationally defined tool. Study 10 was introduced to gain insight into how this form of tool could be developed and implemented.

## 8.5 **Reconsidering the effectiveness criteria**

The original effectiveness criteria (Table 2.3) needed to be reconsidered due to the significant changes in moving from guidelines to design tools. Effectiveness criteria, particularly specificity, had to be re-considered as the design tools had now been deliberately tailored towards designers producing and controlling the quality of design data produced. This was not the case with design guidelines.



The *content emphasis* of guidelines had been an important criterion. The authors of guidelines make their own decisions about how to select and interpret design data to form guidelines. In making judgements about content emphasis, the validity of the guidelines (interpretation of the meaning and significance of a guideline) was dependent on how well the guidelines were written as well on the designer's interpretation of them. As design data were now generated within the design process, content emphasis would be controlled by the designer rather than by an external author. Therefore *validity* of these data was now a concern and dependent on internal factors, such as:

- *Design and management of design tool exercises* - this includes tasks such as: the design of cards; selection and representation of objects and tasks; and the selection of workshop goals and motives.
- *Procedural control of the design tools* - the ability to: refine or adapt design tools to alter design data outputs; manage collaborative discussion between designers and participants
- *Quality of experiential knowledge* - how this is recognised and used.
- *Interpretation of design data* - effectiveness in interpreting data at different levels of abstraction with adequate depth and breadth

Furthermore, as the design tools increase the control and management of design data internally produced, the reliability of these design data could be reduced (consistency in resolving the same type and number of interaction design issues across different design problems or workshops). Reliability could be affected by the following factors:

- *Design and management of design tool exercises* - this includes tasks such as: accuracy of verbal and written instructions given to the participants; recording of design tool procedures and exercises during and between



their use; and choice and selection of participants and designers between design tools and/or workshops

- *Changes in procedural understanding of the design tools* - changes in skills to implement and control design tools by participants and designers between design tools and/or workshops
- *Consistency of experiential knowledge* - depth and breadth of designers and participants knowledge used between design tools and/or workshops

The term *validity and reliability of design data* became a 'specificity', rather than an 'applicability' factor and therefore replaced *content emphasis* as an evaluation criterion. This more accurately described how appropriate, relevant, and effective data were contingent on setting clear, internally derived objectives and procedures and not dependent on having a good *content emphasis* between a design problem and the use of guidelines.

The *structure and presentation* of guidelines was also no longer a relevant criterion. The intention was that design guidance, perhaps in the form of user requirements and interface design guidelines, would be formulated through the development and evaluation of interaction models. The effectiveness of this new form of 'design guidance' would be dependent upon the quality of interpretation from the interaction design models. The emphasis had shifted to the *interpretation of interaction models* to produce user requirements, design solutions and interface design guidelines, however these might manifest themselves.

The proposed design tools still needed to be evaluated in terms of the scope of usability problems they were capable of identifying. The design tools used cognitive dimensions rather than ergonomics criteria to do this, as these appeared to be more accommodating for novel interaction styles. Therefore this criterion was broadened to embrace other taxonomies and so phrased as the *scope of usability issues*.



The other criteria remained unchanged; a summary of the revised criteria is included in Table 8.2

Specificity Criteria	Applicability Criteria
Reliability and validity of design data - Capture and use of data, scope, extensiveness, accuracy and relevance to produce interaction design models	Usability of design tool - level of training and instruction required for designers and participants, level of on-going support required
Interpretation of interaction model - ease with which design data can be converted into a model, limits of understanding, transparency, effectiveness with which model assists in producing design guidelines	Relevance to proposed target audience - level of acceptance and uptake, quality of team working  Likelihood of survival or usage in organisational context - level of acceptance and uptake
Scope of usability issues identified - are all usability issues addressed?	
Supporting novel interaction styles - Ability to address novel interaction styles competently	

Table 8.2 Revised effectiveness criteria

8.6 Study 9 - Evaluation of the card sorting design tool

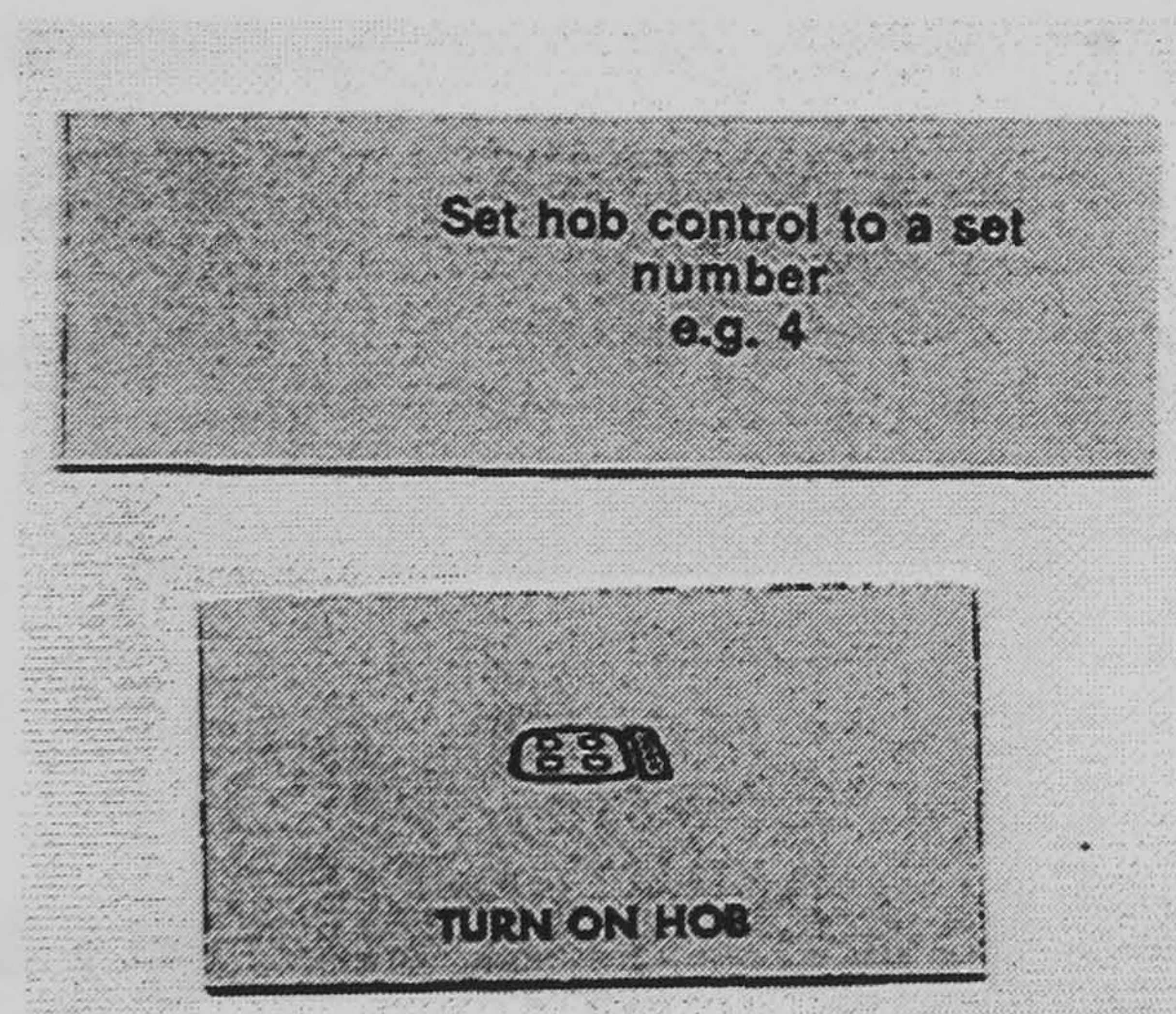
This study was limited to the evaluation of the card sorting tool. This was to allow an understanding of how such a tool might be implemented, refined and accepted by the designers and Design Manager at Electrolux. Some reservations had already been expressed by the design team about carrying out design activity in collaboration with potential end users while having an apparent lack of control over design outcomes.

The study involved three UK industrial designers and a Swedish industrial designer who was currently based in Denmark. As a secondary objective, it provided an opportunity for the Swedish designer to identify new advanced cooker features for an interface design project he was working on.



### 8.6.1 Evaluation plan

The study was conducted over a two-day period. The researcher selected a scenario; this was the preparation and cooking of a Sunday meal. A set of speculative cards was prepared for the design team to reduce their time commitment to the study. These included a wide range of cards that would support all four exercises (as described in Section 8.4.1). Two designers participated in a pilot study. The designers 'walked through' the prepared exercises, sorting the cards while acting as both designers and participants. This was to help familiarise themselves with the exercises and to look for anticipated problems or potential misunderstandings that might arise from the card designs. During this walkthrough the text or legends on each card were improved by supplementing them with graphic symbols. Card statements were accompanied with a symbol or icon representing similar or related task activities. For example, a card depicting the cooker hob was made more meaningful by including a 'hob' image (see Figure 8.2).



**Figure 8.2 Example of original (top) and revised (bottom) cards**

For the main study, four groups of between 5 - 6 participants were recruited for workshops sessions to last about 2 hours. The participants were recruited by the designer team and consisted of factory and office volunteer workers from one of the manufacturing plants. Workshops groups contained a mixture of male and female participants. No attempt was made to assign participants to any particular group. Each group was given some



introductory explanation by the researcher who remained present throughout all workshops. This included the purpose, to prepare and cook a meal and explain the process by placing cards on a table, and an outline description of the tools. They were asked to discuss the process as a group and arrive at consensus agreement if any differences in opinion were found. They were also told that the main objective was to evaluate the design tool and not their performance. Workshop sessions were video recorded. For each exercise, participants were given specific instructions on how to carry out the task. A time limit of 15 minutes was set on each exercise. All groups were given the four exercises to complete, which were presented in the same order to each group.

After the exercises, participants discussed their thoughts on personal cooking habits, perceptions of technology and the effectiveness of the card sorting exercises. After the participants had left, the designers then offered their opinions on the card sorting exercises. These were recorded on video. Notes were made during the workshops and the videotapes were analysed using the evaluation criteria described above. The tapes were then subsequently reviewed. Key events and observations were then made against the revised effectiveness criteria.

### **8.6.2 Observations**

All four workshops had a similar procedural flow. The designers remained generally passive throughout the card sorting exercises and only intervened towards the end of each workshop when more informal discussions began. In all groups, a leader or chair emerged acting as the 'controller' of the cards and also of group decision making strategies. The assumed leader often re-evaluated their task plan to ensure coherence to procedural descriptions.

#### *Reliability and validity of design data*

The design data (evidence and information generated from the workshops that could influence design decisions) generated from the workshops were



rich and variable, but the management and control of design data were negligible, thus resulting in poor validity. The exercises provoked wide and interesting anecdotal discussions about cooking methods, preferences and strategies for using cooking technology, but was not controlled or steered by the designers. This was understandable at this stage. Few facilitating procedures had been given to the designers prior to the workshops due to the uncertainty about their involvement in using the design tool. The participants gathered contextual data, but not in a controlled or predictable manner. For example, there was a tendency to add 'peripheral' or unimportant task or event cards to the task plan, like adding more utensils, rather than adding any real new tasks that provided deeper design insights.

The reliability of the design tool management was high. Participant behaviour across all four workshops was generally consistent. In all workshops considerable anecdotal cooking habits were discussed, prompted by the card sorting activities. Different sorting strategies were adopted within each group, but comments made by the participants and the quality of design data were similar across all groups. Decisions were less dependent upon the exercise task and based more on broad collective experiences thus affecting reliability. Comments were frequently based on family habits '*if we warm food up we use the microwave*' rather than being dependent on the card sorting exercises.

During all workshops, some discussion was given over to reviewing the task plan after a natural phase of cooking activity had been discussed or a reasonable 'component' of activity had been described through the cards. Omissions in task elements were identified through this process and improved the quality of capturing design data. This checking procedure revealed how cognitively different card sorting activity was to real cooking tasks. Participants needed to remind themselves of procedural steps and ensure these were accurately reflected in the task plan, as one participant reflected, '*this is harder than doing the real thing*'. The cards forced participants



to deconstruct activity but not necessarily in a natural manner. In the example below, Participant 1 (P1) was confused whether a completed task (Potatoes cooked) had been represented.

---

P1: 'you've turned them on...to put the potatoes on...didn't you?' [pointing to 'Potatoes Cooked' card]  
 P2: 'Ah but the..'  
 P3: ' you don't need to put them on the hob yet'  
 P1: 'should I put them in?' [the oven]  
 P2: 'go then yes....yes'  
 P3: 'you've got to prepare your veg now'

---

The task plan did not naturally suggest where sub tasks start or finish. Participants found it difficult to model time-related conditions that would be obvious during real activity.

During the second exercise, inserting function cards into the task plan from the first exercise, participants took a less purposeful or contemplative approach. Function cards were read out and participants arbitrarily inserted them into the task map without consideration to their importance or frequency of use. This would affect the reliability of experiential knowledge used in the workshops.

---

P1: 'Do you want an electric helper?' [reading from card]  
 P2: 'that would be at the beginning [of the task map] you'd want to know how to boil an egg'  
 P1: 'Would you want that to be electric?'  
 P2: 'Yer - she wants er...'  
 P3: 'I need help when I cook'  
 P1: 'Does it have to be electric or a book?'  
 P2: 'A book'  
 P3: 'A book might be easier to use'  
 P1: 'Why - haven't you read it? [Laughs]  
 P2: 'Do you mean an electric helper?' [Suggestion]  
 P1: 'Would you use it?'  
 P3: 'Yes I would'  
 D1: 'What about an electronic book?' [Suggestion - but related to one of the function cards provided]  
 P3: 'That would be even better'  
 D2: 'This might be an electronic note book'

---

This exercise, however, did prompt discussions around the activity of cooking, leading to more receptive discussions on the use of technology to support the cooking of new or unusual meals. In the dialogue above, a



subgroup (P1-3) discussed how and when they might use some form of computer-based cooking assistant. They were unclear how it might function or how it might be used. The designer (D1) offered a more focussed solution but did not probe further. During one discussion of this nature, one designer was surprised to observe participants making contradictory demands for technology. Electrolux used four categories in Marketing to describe purchasers or users of cookers. Some of the participants appeared initially to comply with one of these categories, 'Turtles' (weary of technology) but later on were happy to consider quite radical and advanced proposals thus demonstrating behaviour usually associated with a different user Marketing category known as 'Hedonists'.

The third and fourth exercises were less successful in producing design data. In the 'week-in-the-life' exercise, participants generally added typical meal types under each week-day heading without discussion. The final exercise, 'character profiles' generated stereotypical comments about the usage of technology and again did little to reveal any insightful comments that could be effectively used as design data.

After the workshops, designers were asked to comment on the effectiveness of the design tool. Comments were very favourable but there was little evidence, apart from the task map, that other forms of design data had been captured. After the workshops had been completed, a summary statement was collaboratively drawn up with the research investigator.

The statement below reflected a user-centred tone with a strong emphasis on perception towards technology rather than on specific functional requirements. This was very encouraging. It was hoped that this type of design data would be achieved from the design tool but further work was required on improving the reliability and validity of the design data.



---

"It is important that the cooker interface instils trust to the user by providing ample feedback and information on the consequences of using any new or novel technology. New features will not be used unless the user fully understands the implications of such an action and can be confident that the action has been accepted by the cooker.

The interface should avoid providing functions that are 'owned' by the cooker rather than the user, for example timing devices where the cooker is allowed to own some time keeping tasks. Controls should always suggest that the user is in charge by permitting clear and positive feedback of their purpose but should also allow more adventurous users to feel 'master' of the cooker by allowing some controls to be configured to their own needs.

The interface should provide a 'supporting' rather than 'expert' role either in terms of food safety and hygiene or in introducing the user to new methods of cooking or new types of cuisine.

The key criterion, however, by which any novel features for a proposed interface must be assessed is on TRUST."

---

### *Effectiveness of interaction models*

The task plan was intended to form the backbone of the interaction model with each exercise building up a contextually oriented representation of cooking activity. The plans, card settings and comments produced by the participants were intended as a record of design data. The construction of the task plans was video recorded. However, the designers paid little attention either to their construction or the completed plans. This appeared, at the time, to be disappointing. Design data, in the form of the task map, was not formally documented but remained personal reactions and experiences. Little was shared between the designers about the knowledge gained. Indeed some of the designers had come to the workshops with a personal design agenda and saw the workshops as an opportunity to test their ideas.

Another purpose of the model was to create a common dialogue between designers and participants; clearly this did not always occur. In the example below, the designer (D) tried to identify why P had such little faith in the safety of her electrical products. Terms like 'hot product' are commonly used in Electrolux but unfamiliar to the participant and discussion in not pursued because of this. Many of these deficiencies were attributed to a lack of shared understanding between the designers and participants. The design tools need to ensure that this form of breakdown does not occur.



---

D: 'You mean you re-set the clock every day?'  
P: 'Me electric goes off at the wall [at night time]. I could not afford for me house to be burnt down, for safety everything goes off at the wall'  
D: 'Is this because this is a "hot product?"  
P: 'No - it's all off at the wall'  
D: 'Everything?' [meaning other electrical products]  
P: 'Every electric product'

---

### *Scope of usability issues identified*

An interesting ergonomics issue emerged from the workshops. This was to do with trusting technology and became an important determining criterion by which technology was assessed. The workshops provided an opportunity for less obvious ergonomics factors to be discussed. This suggests that acceptance criteria would also need to be part of any future proposed 'scoping' criteria. Participants perceived technology as belonging to 'us' or 'them'. One group presented themselves as collectively not having an interest in technology but could see that other demographic groups might be interested.

---

'me Brother does, but he's 15 years younger - he does. [use and understand technology] It's just our age [refers to other members of the group] we don't ..er.. know about it.

---

### *Support novel interactions styles*

In designing the card sorting activities, it was difficult to understand how novel concepts could be introduced or how they might be generated during the exercises. One approach emerged entirely by accident. Writing vague or ambiguous statements on the cards often prompted questions about their meaning or significance within the card sorting exercises. While clarifying their meaning, suggestions were often put forward, which occasionally resulted in creative proposals.

### *Usability of design tool*

Assessing effectiveness in terms of usability depends upon what is being measured. There was no doubt that the cards proved an effective vehicle for



promoting discussion. The exercises themselves were not difficult to accomplish, although procedural problems were identified. For example, participants often needed reassurance on 'rules of the game'. Task planning exercises were very time consuming and often had to be reviewed for consistency and errors.

For the designers, the level of engagement with the design tool was very low. They did not intervene in the card sorting activities and only occasionally offered advice. When asked why they had not taken notes for use later on, they stated that they did not feel it was necessary as the process had provided them with many new ideas. They felt they had a clear understanding of the direction they would take with future cooker interface proposals.

#### *Relevance to proposed target audience*

The designers were very encouraged by the workshops and found the exercises extremely illuminating and worthwhile. One designer said '*in the five years I've been here I have never been able to gather as much useful information from users as I've been able to do here*'. The design tool provided an opportunity to involve users in a collaborative rather than consultative role. Persuading the design team to embark on such a process was, at times, difficult and required a great deal of 'hand-holding' from the researcher. Many of the procedural elements of the design tool were untested which contributed to their sense of unease about using them. Due to this uncertainty, it was difficult to assign the designers clear roles resulting in them becoming passive observers. Shifting ownership of the design tools to the designers became the next important iterative step.

#### *Survival of design tool*

Prior to the workshops, all designers expressed concern about the card sorting design tool and were hesitant about a process they were not directly in control of. They were unsure how participants would react to vague or



unclear proposals and did not relish the prospect of deliberately placing themselves in a situation where they had uncertain or no design proposals to offer the participants. As one of the designers said, *'we don't want them going away thinking we can't design a cooker'*. Certainly the designers felt no ownership of the design tool before the workshops. However, afterwards the reaction was very different. They were encouraged with how participants dealt with the situation and surprised at the participants' level of creativity.

---

'It was a nice session, and I believe people got something to think about. I divided the team into three groups. "Hedonist", "Committed" and "Nostalgic". We have some information how these type of consumers are acting. I didn't follow your description, I skipped the last two exercises. I was short of time. I asked them to evaluate the features so it ended up with drawings on what they believed was an ideal cooker. It's difficult to evaluate how much impact the preparing part has (referring here to the first exercise - task planning). Let say that it could be interesting to find out what kind of raw material they buy, how they store it etc. I do believe it's interesting, but how to judge that in relation to the design of timers, knob etc I don't know.'

---

One designer was sufficiently enthused that he decided to carry out his own workshop independently in Denmark. The designer provided some feedback on this workshop. He used the second card sorting exercises to test some radical design features but found deriving focussed design data difficult. More reliable and valid methods of capturing design data needed to be investigated.

### 8.6.3 Reflections

The validity of the design data produced from the design tools still required improvement. Exercises needed to be set against specific design issues or problems currently being investigated within the design team to be effective. The relationship between objectives and outcomes of the card sorting exercises needed to be more clearly defined and be explicit for both the designers and participants. Although a rich source of design data was gathered, the designers were not equipped to capture or control the type and quality of design data generated. Designers were unclear of their role.



Nevertheless, there was sufficient evidence to suggest that contextually based design data could be gathered.

More guidance was required on designing the cards for example, coding methods, illustrations, colour, and shape of cards. Training exercises for both the designers and the participants was considered a possibility. Card composition required more consideration to accurately trigger discussion about potential user behaviour and needs.

The 'interaction model' was not used effectively. Inserting function cards at recognised stages rather than being considered globally or by forcing participants to consider the implications of adding functions into the task map may have helped participants to consider the implications more deeply.

It has to be recognised that the quality and detail of the design data using this type of secondary design tool will be 'attenuated'. The quality and modelling will inevitably be less detailed than using ethnographic or participative studies using trained designers or researchers. Subsequent studies were set up to establish what kind of balance needed to be set between attenuation of data gathering and the effectiveness of using these data.

### **8.7 Study 10 - Student user trials with design tools**

A further phase of design tool development was planned using all three design tools in an integrated fashion. However, this was not possible due to some minor re-organisation at the UK and Danish design studios. It was unlikely that any members of the design team, used in this study, would be able to continue their support for the subsequent months. To continue the development of the design tools, an alternative evaluation plan had to be implemented. An opportunity arose where industrial design students could be introduced to the design tools.



Observing student designers presented a dilemma in evaluation planning. Students have very different motivations to practising designers and do not have access or control over potential participants in the same way. Design problems are also contrived to deal with student assessment. Student projects do not reflect live design projects by having to contend with social and commercial 'noise'. For these reasons, evaluation objectives from this type of study had to be more specific and controlled. It was not possible to use the effectiveness criteria in the same way as the previous study. However, this study could address the issue of ownership by allowing student designers to tailor the design tools to their needs. It was anticipated that this would help identify factors that would make design tools acceptable to (student) designers.

Two evaluation studies were carried out with graduate and under-graduate students thus creating two iterations of action research. Both iterations involved students from Teesside University. In the first iteration, six post-graduate MA Digital Design students on an optional 15-week module in 'Practice in interaction design' were involved. In the second iteration, nineteen second-year industrial design (BA Hons) students undertook an interface design project as part of a 5-week core module.

### **8.7.1 *Evaluation plan for MA students***

Students were aware that some of their studies would involve untried methodologies and therefore clear guidance would not always be possible. Five of the students had degrees in industrial design and one in fine art. All students worked on their own interface design proposals, these were:

- personal library support aid or assistant;
- interface for a 'Quick-time' virtual reality camera;
- digital audio mixing unit;
- digital photo album;
- teaching aid on the digestive system for primary school children;



- digital tourist map and navigational unit.

Each student worked independently and project work was conducted on each design tool consecutively. Opportunity was given to reflect on each design tool before embarking on the next through tutorial sessions. On completion of this work, students were asked to provide a written methodological explanation, their findings and a critique of the exercises.

Students were first given a tutorial on the card sorting tool with practical demonstrations and walkthrough sessions. Procedural problems identified in Study 9 were presented to them and possible solutions were discussed. A second tutorial was provided on the inspection-based tool to explain the cognitive dimensions and some advice was given on how they could be used within their proposed projects.

The scenario design tool was the least defined tool and therefore more open to interpretation. In order for the students to understand the rationale behind this design tool, a field-based workshop was carried out with one of the student's product proposals, (tourist map and navigational unit). The student was asked to prepare a block model of the device and be able to describe, in high level terms, the functionality of the device and how a user might interact with it. The researcher and all the students were involved in this workshop.

A scenario was then defined, where a tourist needed to find their way from a railway station to a Tourist Information Centre. A role playing scenario was then carried out where one of the students acted the role of 'tourist' at Middlesbrough railway station using the 'prototype' device while the 'designer' talked through the features and functions as they were requested by the 'tourist'. This provided an opportunity to explore how the scenario design tool could be developed so that it might adequately support the designer through such experience.



### 8.7.2 Observations

During discussions with the MA students, a possible solution emerged to preventing participants blindly adding functional cards without prior consideration. The notion of a *function filter* emerged. Before inserting function cards into the task plan, each card was assessed against two criteria, possible frequency of use and possible importance to the task. A 2x2 matrix was created, 'frequently used' / 'not frequently' used along one axis and 'important' and 'not important' along the other. Cards entered into the 'frequently used and important' cell would then be inserted into the task map first. Students took up this suggestion for their own card sorting exercises.

Due to the ineffectiveness of the third and fourth exercises in Study 9, these were removed. They conducted their own card sorting exercises with at least 3-4 potential users for their proposed product. The labelling of cards created the most problems. Students found it difficult to produce cards that would prompt creative dialogue about future or novel interaction styles. Students commented that their cards depicted very specific task actions and were not open ended enough in their interpretation to encourage speculative discussion. One of the students used card sorting on children and reported a high level of acceptance by the children although he observed a heavy degree of peer pressure in arriving at conclusions. In common with the previous study, students reported surprise at the level of insight some of their participants demonstrated.

While developing the scenario design tool, there was a shift in 'control' between the student 'tourist' and student 'designer'. The student 'designer' began very confidently by walking through the features as required by the 'tourist'. However, as the scenario design activity progressed, the student 'tourist' made more detailed and challenging demands on the exact functionality of the device. Interaction design issues were emerging beyond the 'student-designer's' original problem frame. For example, one important design problem that emerged was how display-maps should be oriented?



Should maps be fixed to the device or the environment? An issue the designer had not thought of. This exercise provided the MA students with an opportunity to understand the need for careful planning for the scenario design tool. To deal with this issue, the students decided to conduct some scenario design work between themselves before working independently. However, this was carried out in a design studio as a group discussion and not within the context of a real situated environment. Unfortunately this proved unproductive and the students reported that most of their discussions centred on re-evaluating their own design proposals rather than improving the effectiveness of the scenario design tool. The students therefore carried on with their own scenario design activity without any clear procedural rules.

The MA students provided written feedback on their views and perception of the design tools. Most of them were extremely complimentary of the tools and used them in subsequent design work in other projects. Many of them made comments on how it had affected their way of thinking about their role as a designer. The students expressed surprise at the quality and inventiveness of suggestions made by participants and how productive design solutions emerged from collaborative discussions with potential end-users and demonstrated more willingness to shift the locus of design control to users. As one student commented, it '*ensured that it was the users making the decisions and not me*'. Generally, they thought the card sorting tool worked well as a high-level concept generating and defining tool and that the scenario design tool encouraged convergent interaction design solutions. Although there was still concern about how best to control this convergence in order that a single solution was derived. Due to their inexperience, they resolved this problem by suggesting their own design solutions for participants to approve rather than exploring the design problem in more depth. The quality of feedback was reported to be variable and dependent on the participants selected. Some noted how a dominant participant could affect other participant's opinions. There was a recognition of the benefits of



presenting interaction design problems within a situated context, *'people involved in a scenario means that it will have a good grounding in reality'* thus providing opportunities for *'ideas to flow'*.

The most common criticism of the design tools was placing the inspection-based tool between card sorting and scenario design. This forced the designers to address interaction design problems at a detailed level too quickly, and it was *'trying to lay down too many rules at a too early stage of the process'*. The students found them very difficult to use. Despite explanations given in a tutorial, they still found the vocabulary difficult to grasp and difficult to define an abstraction level in which they could use the cognitive dimensions. They were unclear whether each dimension should be judged against the whole interface proposal or at lower levels of abstraction. There was also confusion about the relationship between each dimension and priorities between them and one student reported that he found them repetitive to use. However, another student thought they offered a good filtering process and provided structured thinking.

Suggestions were offered about how the integration of the design tools could be improved. The design tools were taught, implemented and evaluated at different points in the module, forcing their use to be fragmented. Reducing time intervals between the design tools was considered a useful improvement so that design data was kept more *'alive'*. More support was also thought important in defining, managing and controlling scenarios and card sorting, for example, *'I think I would have got more from this part of the process [card sorting] if I knew exactly what level of detail would get the best results'*. One student reported the difficulty of providing a balance between too much complete information and providing too little. A complete, fully detailed prototype may reduce the level of creative options while a prototype too low in detail may not suggest sufficient design boundaries to allow useful solutions to emerge. Students suggested that each tool had different strengths and weaknesses and used in combination helped develop a deeper



understanding and framing of their design problems. Through continued use, they thought, the tools could easily be adapted and refined to suit their needs.

8.7.3 Changes made to BA students' evaluation

The first two card sorting exercises remained the same as in the MA student module, and included the additional 'filter' exercise. For the cognitive dimensions tool, a matrix was developed to reduce the problem of devising levels of analysis.

Cognitive Dimension	Action							
Abstraction Gradient	Entering Data	Reading Data	Checking and confirming	Adjusting controls	Adjusting displays	Selecting functions	Cancelling functions	Obtaining help
Steep	x			x	x		x	
Slight		x	x			x		x
Comments				To start with, but familiarity with system makes it easier	Becomes slight once displays are chosen		On/off could be used as a cancel button	System takes you through step-by-step

Table 8.3 Examples of cognitive dimensions structured around quantitative values

The cognitive dimensions were listed on the vertical axis and a range of interaction sub-tasks were described along the horizontal axis such as 'entering data', 'checking and confirming' and 'selecting functions'. Within each cell dimensions could be evaluated against a specific sub-task. Figure 8.3 illustrates the dimension of 'abstraction gradient' with some sample annotation from one of the students.

8.7.4 BA student evaluation plan

The 19 undergraduate students were divided into groups with 3 groups of five students and one group of 4 students. Students were provided with a group-based brief to design a novel interface for a microwave oven or a portable information device (see Figure 8.3), both specifically designed for students. They were instructed that they were not constrained by



conventional control and display technology but any final solution must conform to usability criteria developed from the design tools. Students verbally presented their work in groups and submitted their group design work for assessment.

#### 8.7.5 Observations

Card sorting was the most popular activity and most students reported obtaining useful design data, although many of the groups reported participant fatigue. Two of the groups used the scenario design tool with considerable imagination and flair and found methods of overcoming many of the problems identified in the MA study. One group improved the judgement and handling of design concepts by developing a selection board where participants could choose function and/or interaction styles to suit their needs at a particular point in a given scenario. Different types of control and display options were offered as a series of simple sketches and annotations. When a participant required a particular function, it could be chosen from a range of variant design solutions or options represented on post-it notes. In this way participants could quickly build up an interface from component elements. This idea was carried forward to future studies and became known as the *tab board*.

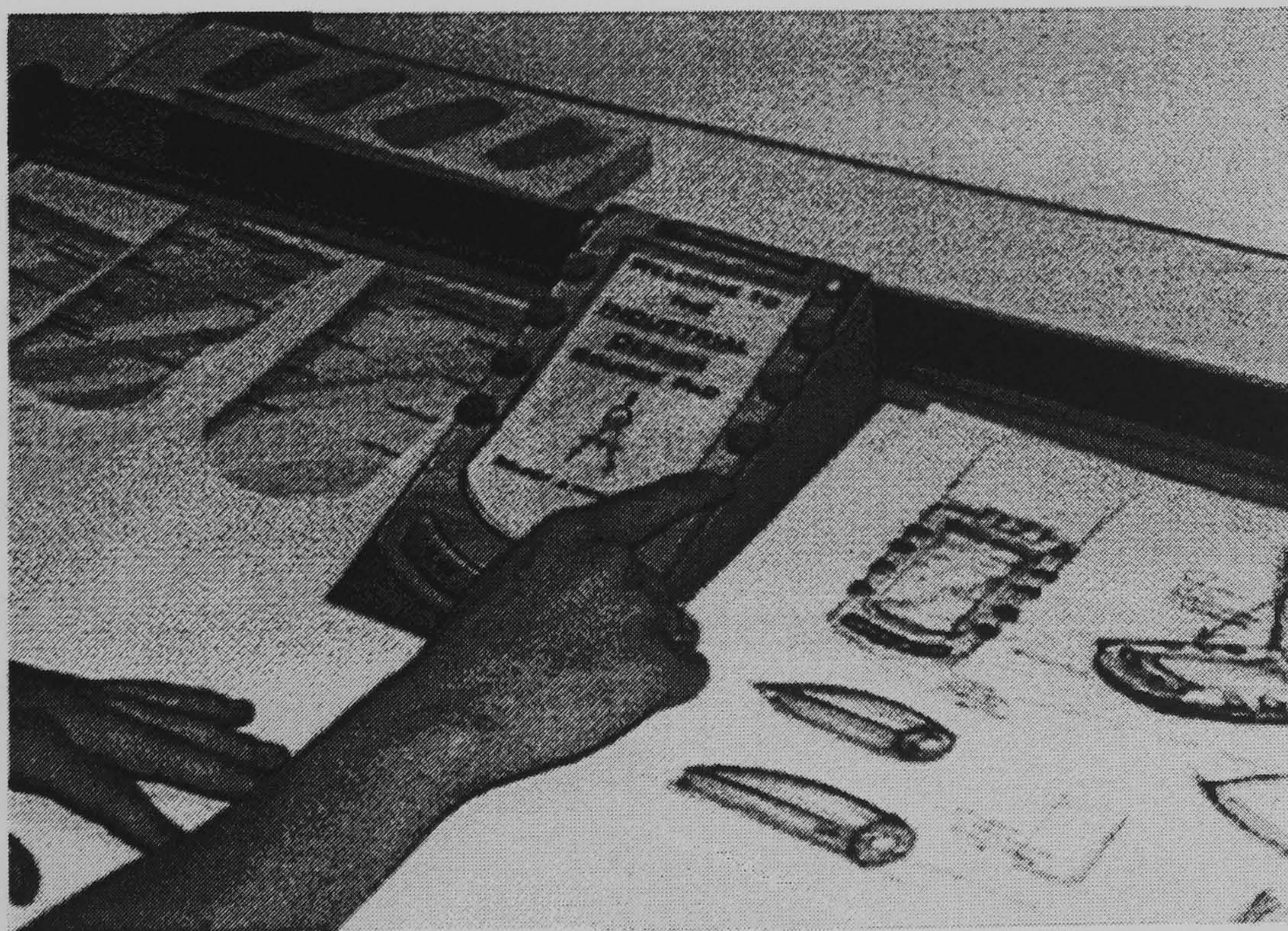




Figure 8.3 Example of student project

Many of the other groups found it difficult to manage role playing with participants who were also their own peers. Students still had problems in using the improved inspection method and generally did not use the method effectively. The only positive comment was its usefulness as a checklist to ensure that all interaction tasks had been considered. Overall, the undergraduate students demonstrated less motivation to using the design tools than the MA students did; much of this could be attributed to working in groups.

### 8.7.6 *Reflections*

It was unfortunate the refinement of the design tools could not have been continued with practising designers. The use of student designers had obvious limitations and restricted the scope of development. Design students were motivated by assessment criteria, which inevitably distorted design behaviour. Some degree of caution is required as students are motivated by assessment grades and were less likely to criticise their perceived effectiveness of the design tools. Students had more time, were willing to experiment and could afford to take more risks than the practising designers in Study 9. Therefore it would be unwise to equate student behaviour with commercial practice. Nevertheless, some common traits were identified. Designers, in both this study and Study 9, often adopted a 'solution push' approach. Preconceived design proposals were put forward for consideration more than using the design tools to encourage participants' latent needs. Participants were used in a consultative rather than a collaborative role. The student designers were less apprehensive than the practising designers about engaging with participants and exploring design proposals with them.



This study provided an opportunity for student designers to use and adapt the design tools. Improvements were made, particularly in introducing the functionality filter as an additional interaction model, and the use of 'tabs' to describe device-based components derived from the card sorting tool. The MA students and two BA student groups took ownership of the tools and found the experience of using the tools enlightening.

The inspection tool remained a problem. Due the resistance in accepting such a tool, a more narrowly focussed study was initiated exploring whether *any* form of inspection method was compatible with product design behaviour. A comparative study of three different inspection methods was initiated to explore this issue.

## 8.8 Study 11 - Perceived applicability of three inspection methods

Inspection methods are a useful tool in HCI to quickly identify usability problems without involving user trials. The aim of this study was to understand how industrial and graphic designers, without any formal or informal HCI or human factors training, would *perceive* the usefulness of inspection methods as part of an interface design process. Also, to identify how inspection methods could be improved or adapted to meet their particular needs. Rather than assess the effectiveness of different inspection methods by measuring the number of usability hits identified, the perceived effectiveness of three different methods was measured. This is so the potential uptake, ease of use and effectiveness of such methods could be assessed. Three different methods were compared: heuristic evaluation (HE) (Nielsen 1994), cognitive walkthrough (CW) (Polson *et al*, 1992, Wharton *et al*, 1994) and an adaptation of cognitive dimensions (CD)(Green and Petre 1996). Variants of the cognitive dimensions method had been used in Study 9 and Study 10 (for a description of cognitive dimensions, see Section 8.4.2). Further improvements were made; these included extending the title of each cognitive dimension so that it contained more immediate meaning and a



concise definition. Also, a real world example of an interaction style being defined in terms of a cognitive dimension was provided. Three example cognitive dimensions are presented in Figures 8.2 a, b, and c. The HE and CW methods were selected because of their popularity in many HCI studies and software development projects.

### **8.8.1 Evaluation plan**

Five lecturers from the Institute of Design, Teesside University, UK, who had been practising designers, volunteered to take part in the study. Four lecturers had a background in graphic design and one in industrial design. A repeated measures design was used. Each inspection method was applied to a different novel prototype interface so different usability problems could be found in each condition. The three screen-based novel prototypes evaluated in Study 3 were used. The prototype interfaces all provided the same functionality but used very different interaction styles to perform user actions. All prototype interfaces had known usability problems. None of the designers had any involvement in the design of the prototype interfaces. As none of the methods were known to the subjects, participants were asked to familiarise themselves with each evaluation method using a simple paper based tutorial. A few quick questions were then posed to check their understanding. This was repeated in each condition. Each condition took around 30 minutes to complete and conditions were order balanced. Participants were asked to go through each interface with a different method and talk aloud about their perceptions about using each method. At the end of the study, participants were then asked a set of questions about their preferences for the inspection methods. Interviews were audio taped and later transcribed for analysis. Content analysis (Krippendorff, 1980) was used by dividing the transcripts into units for analysis, which in this case were the smallest stand-alone statements or comments, where some degree of inference could be instantiated. All data from the transcripts were analysed and therefore sampling was not required. Each unit was then analysed for its contextual meaning and categorised thematically. This was



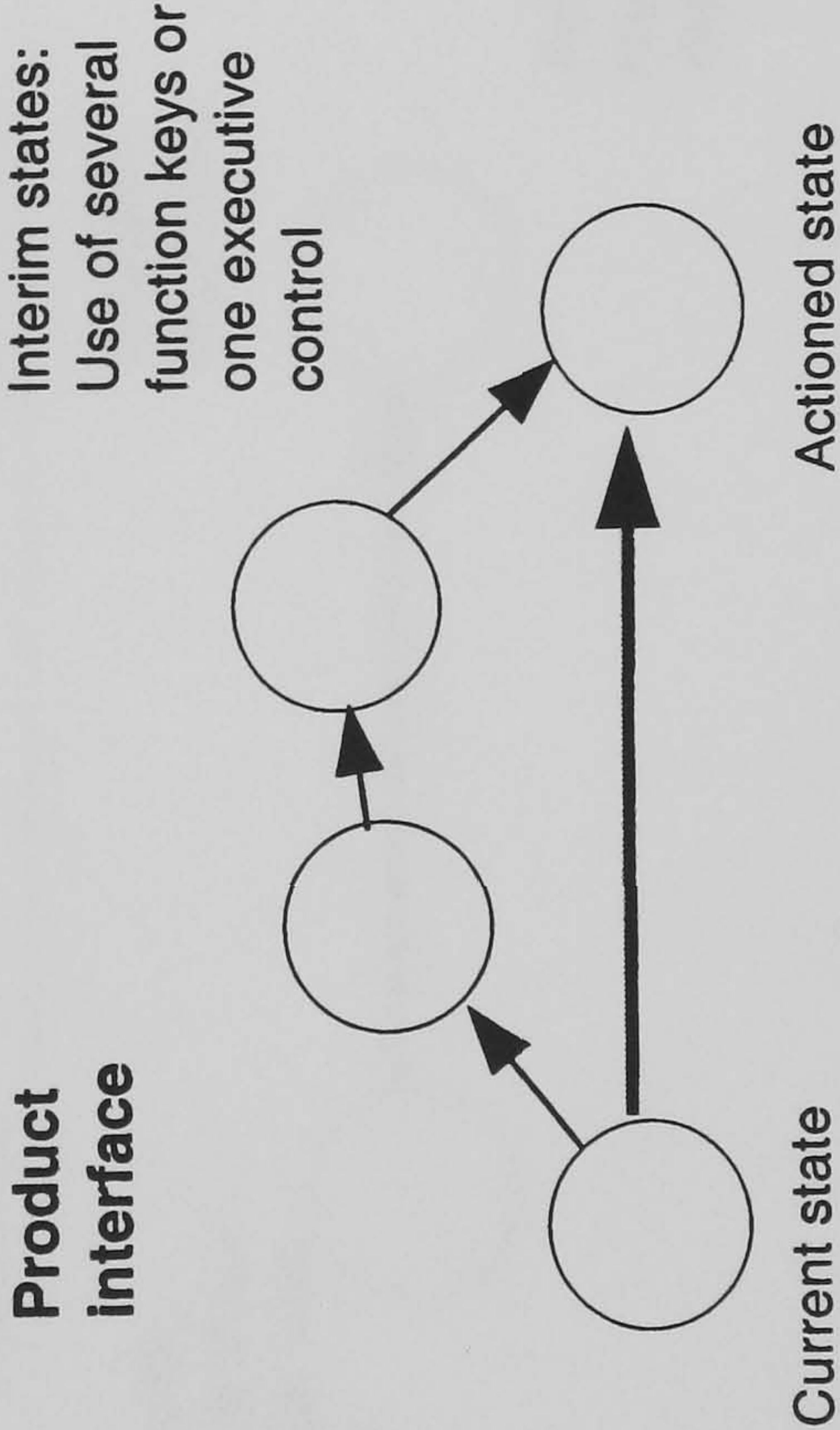
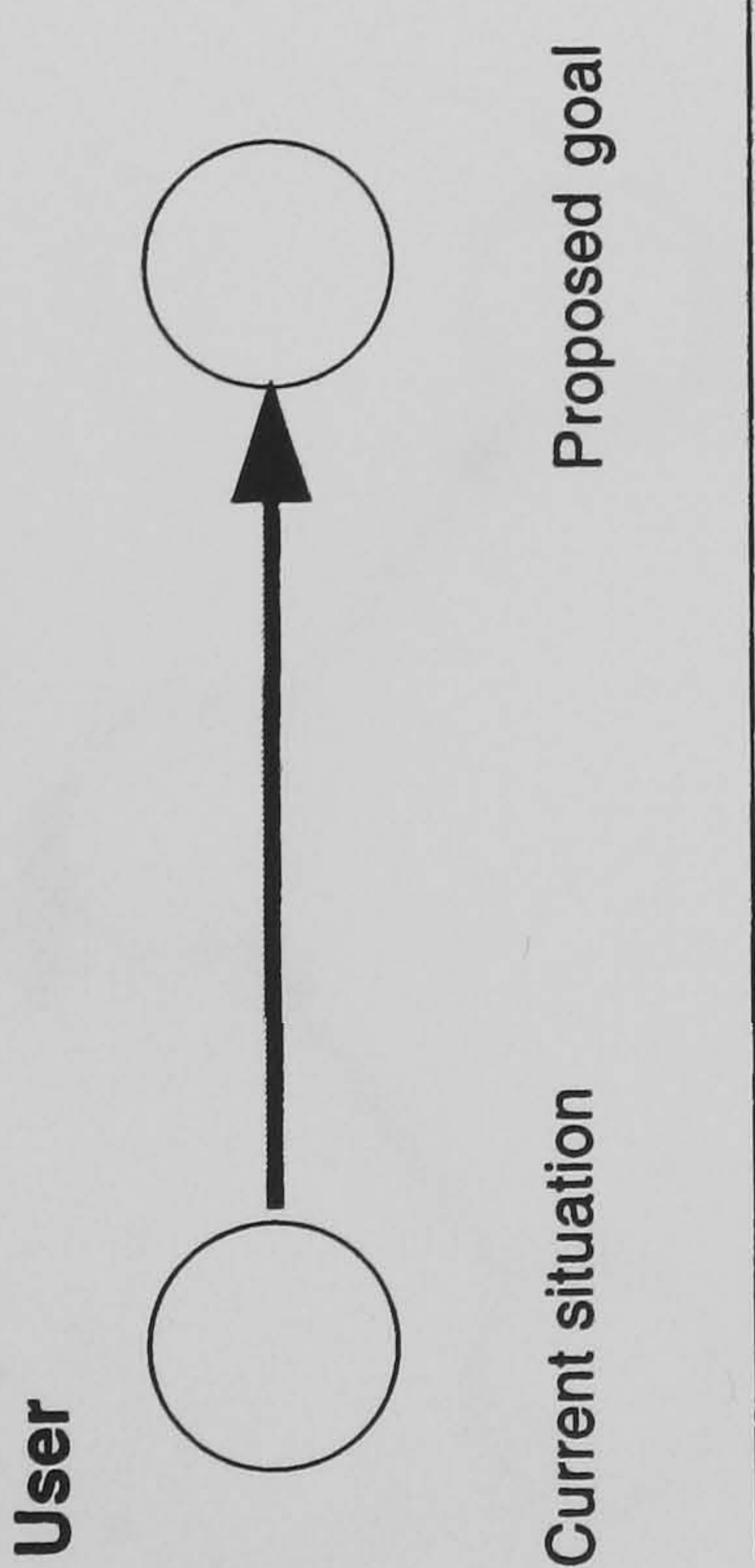
repeated until a consistent and robust taxonomy emerged. Due to the relatively small data set, quantitative analysis was not feasible such as statement or comment frequency counts.



Executive control

This measures the number of steps, actions or procedures that can be grouped together to perform a single 'executive' task, thereby reducing the number of interim states, and also considers the level of difficulty the executive control may impose to arrive at a proposed goal

Diagrammatic explanation



Description

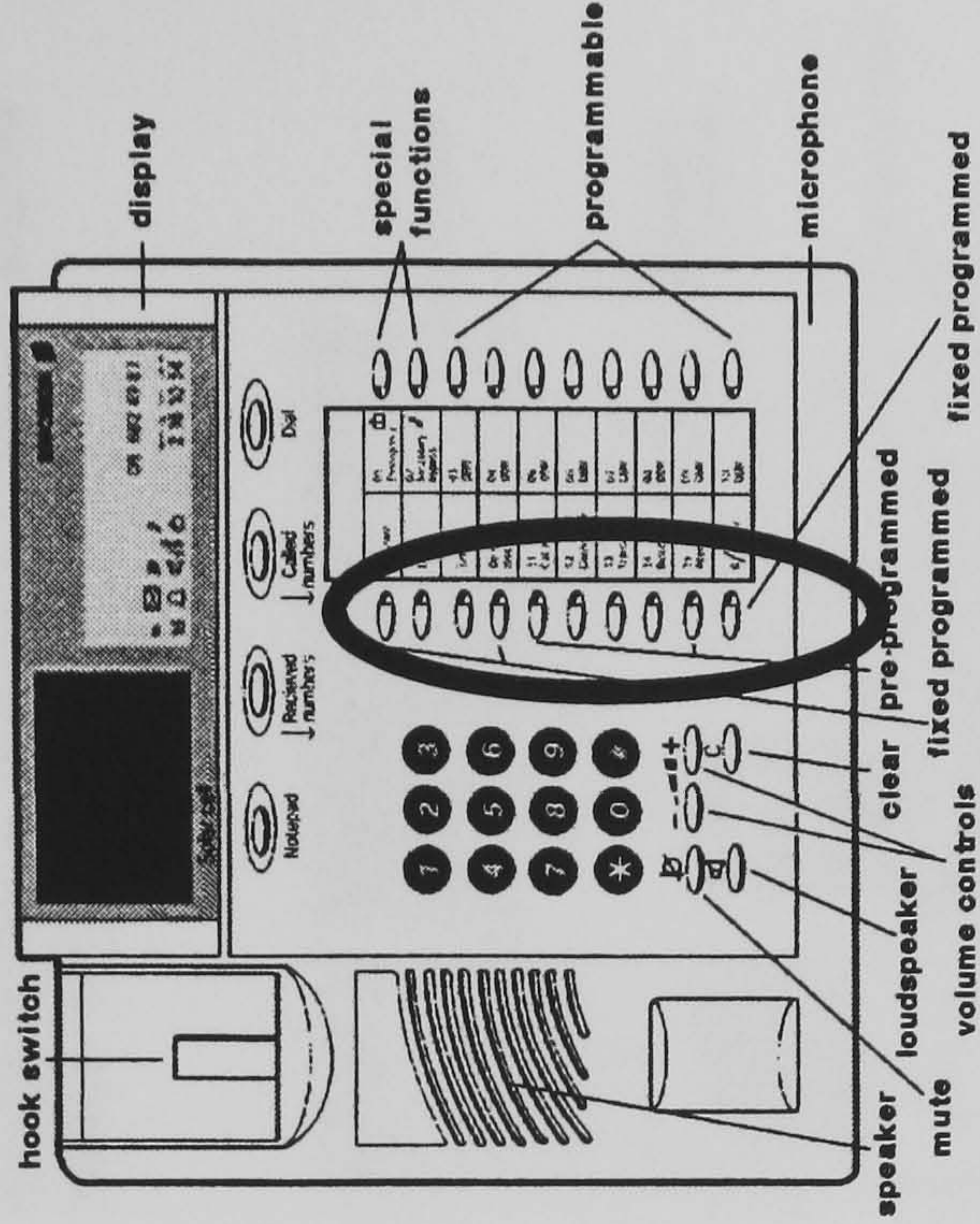
Auto features or memory keys are a good example of executive controls. Some interfaces provide either accessible or inaccessible executive control, in some instances inaccessible executive control can be a good thing, for example where activation of such a control may have disastrous consequences.

In relation to the diagram, the measure of executive control could be expressed in terms of the 'distance' (number of states that are by-passed) or difficulty (how easy it is to activate an executive control) in moving from the current state to the actioned state.

Executive control using good interaction design principles should be:

- appropriate and relevant to typical tasks
- easy to set-up, remember and cancel ideally not requiring the use of a manual
- provide hard access to critical executive controls

Example



Features on this phone, which use executive controls are: transferring from notepad, transferring from received numbers and transferring from called numbers. These programs or executive controls are defined and pre-set by the user to allow quick activation of sequenced-based tasks.

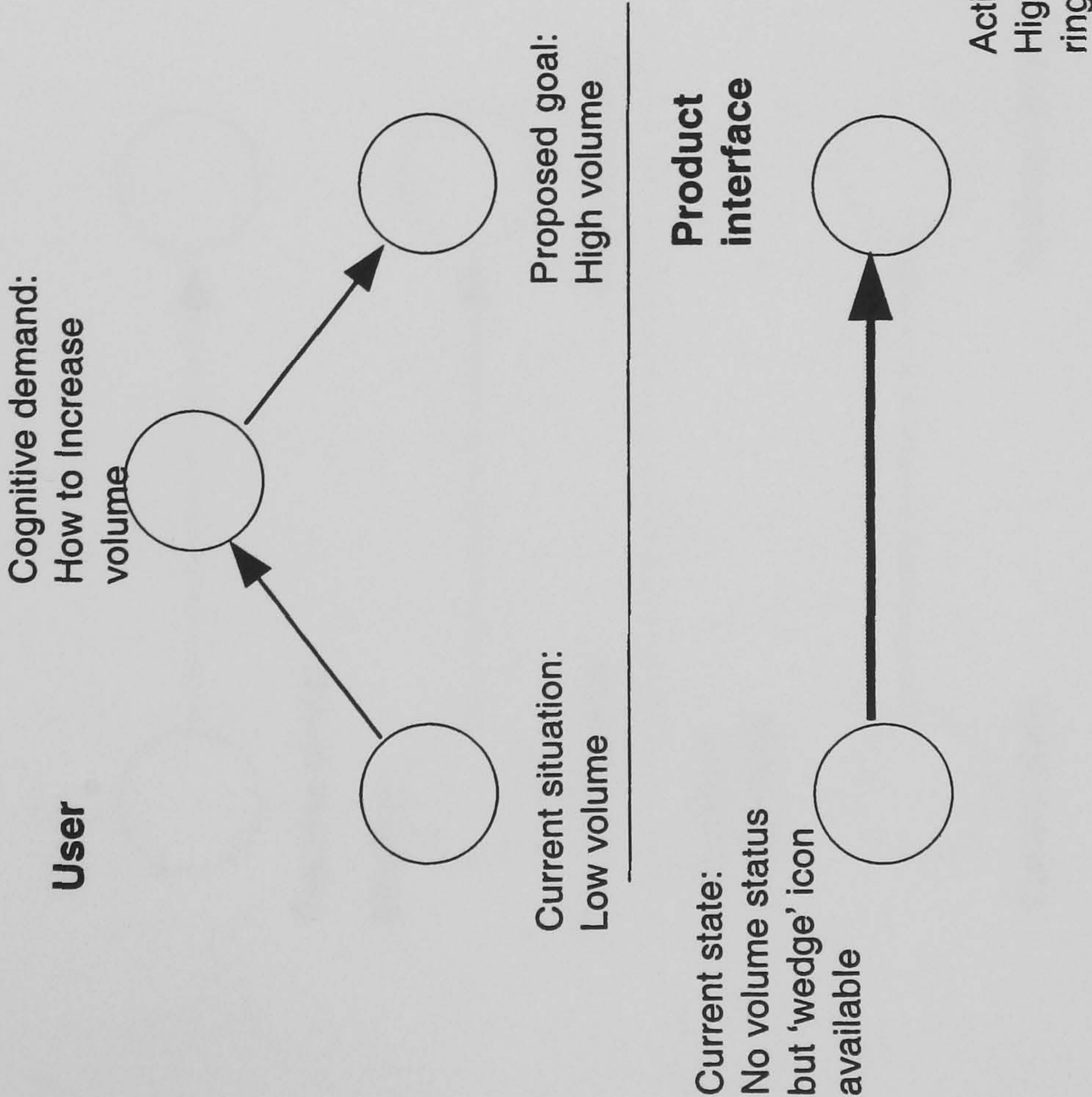
Figure 8.2 (a) Sample Cognitive Dimension



## Closeness of mapping

This dimension measures how well the representation of interface states maps onto the user’s existing ideas and concepts

### Diagrammatic explanation



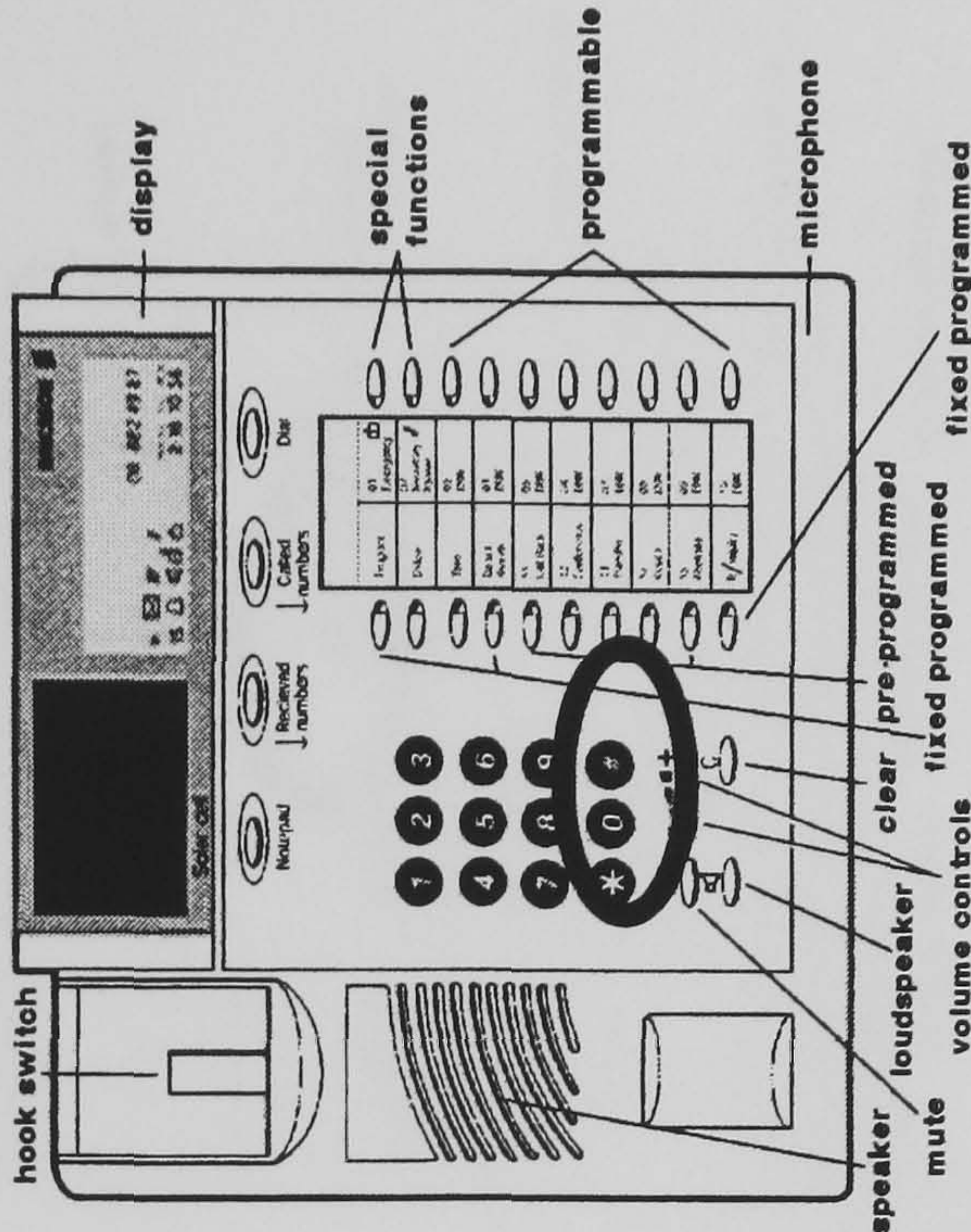
### Description

An interface that fits conceptual ideas and concepts of the user’s tasks or preconceived ideas about the product (for example by using a metaphor) will mean that the user requires low cognitive demand to learn and use the product. A good *fit* between the user’s understanding of the states and their associated task meaning suggests that an interface maps closely.

As a designer, don’t assume you know the best type of analogies or concepts that should be conveyed. Always find these out through the scenario building exercises.

### Example

The volume control uses a mixture of coding to express the concept of increasing and decreasing volume control by the use of ‘-’ and ‘+’ and the ‘wedge’ icon.









### 8.8.2 Observations

It was possible to categorise emerging themes under various headings, these were: ease of learning, perceived 'cost' of using methods, perceived usefulness and general assumptions made about the methods.

#### *Ease of learning*

All methods used unfamiliar language resulting in participants sometimes interpreting or making inferences beyond the intended scope of each method. Few distinctions were made between usability problems and user preferences for example. They were used to identify potential usability problems and also extended to capture perceived subjective impressions or preferences that users potentially might have. Conversely, some usability problems were identified intuitively without using any specific method. One participant stated that he had identified a problem and could not find a way of fitting the problem to the inspection method. Aspects of evaluation methods that were quick to learn or perceived as being adaptable or interpretable were preferred, and participants thought that each method provided a structured checklist-based framework in which to conduct evaluation tasks, although HE was generally considered to be more accessible than the others. However, it was felt that the use of procedural or structural approaches would result in inflexibility or prevent adaptation.

#### *Perceived usefulness*

Participants expressed concern about the scope or coverage of the design tools and were unsure about how successful they had been in identifying usability problems and wanted feedback on their level of success. The designers thought it was important to have control over how and when they might implement these methods so that they were



not 'methodologically' driven approaches being imposed on the design process.

### *Perceived 'cost' of using methods*

The use of evaluation criteria such as breaking a task down into action unit for conformance testing was liked but analysis units were often difficult to define; for example, what was meant by an 'action'.

Participants assumed that the harder the method the more rigorous and comprehensive the usability problems would be. The concept of a high learning overhead was seen as a positive attribute as more insight would be gained once a method had been mastered.

### *Overall benefits*

Overall, participants preferred the HE method for its compactness, accessibility, transparency and adaptability, although concern was expressed that only obvious problems would be uncovered. The method did not provide navigational support or offer a natural 'starting point' from which to begin evaluation. It was also suggested that this method provided stronger hints towards solutions for usability problems.

Participants found the CD method very prescriptive, once the vocabulary had been mastered, but thought that the vocabulary created too much of an inhibiting factor to designers. It was regarded as being procedurally oriented and time consuming, as constant referral to the tables was required. The diagrammatic examples of the dimensions proved very useful in helping to understand their meaning. There was also doubt if the dimensions would identify all usability problems.

The CW method was considered to provide the greatest degree of flexibility and scope in the type of applications it could be applied to but participants expressed concern on the level of dependency required



on the designer's analytical skills. The structured questioning offered a way of uncovering usability problems but, without any formal training in usability evaluation, many usability problems may simply be missed.

#### *General assumptions*

Participants expressed conflicting needs. Formalised procedural evaluation methods were welcomed and regarded as an important conformance test in identifying usability problems and should demonstrate that it has comprehensive scope or coverage in identifying usability problems. However, in contrast, there was a general reluctance to use methods that were perceived as constraining the design process.

#### **8.8.3 Reflections**

Participants were asked to express their views and perceptions of the methods over a very short period of time. In this sense, the manner in which applicability was evaluated was contrived. The learning of the inspection methods was experimentally controlled and not monitored in the field. Nevertheless, this study provided an insight into the methodological *characteristics* that an inspection method must offer to be acceptable to designers without specialist HCI training. To attain good applicability an adapted inspection method should offer adaptability and interpretation to ensure it is methodologically oriented towards design outcomes rather than perceived as imposing an alternative design process.

This study revealed that designers have contradictory demands between formalism and adaptability. These findings are common with Study 5. Attempts were made to overcome this problem by amalgamating the preferred characteristics of the three methods so that it offered a quick accessible checklist with access to the more challenging usability criteria only if required. The adapted method was



inserted into the design tool handbook evaluated in Study 12 (see Appendix 13.5). However, further work is still required measuring the validity of the design data produced from this adapted inspection method.

### 8.9 Conclusions

Studies 9 to 11 contributed towards evolving the proposed design tools from a set of design requirements (see Section 7.4) to a set of integrated secondary design tools through user evaluation. The next step was to evaluate these integrated design tools as part of a live interaction design project. This was carried out in Study 12.



## 9 Final evaluation of design tools

---

### 9.1 Introduction

The refined design tools were procedurally described in a handbook as an accompaniment for the designers involved in Study 12. In this study, issues related to organisational survival were raised prompting a further, and final, investigation (Study 13). This was carried out at another design organisation to explore this issue more deeply. The chapter concludes by reflecting on the findings from both studies.

### 9.2 Aims and objectives

- To evaluate the complete tool set using a live interaction design project against the revised effectiveness criteria

### 9.3 Design tool handbook

The researcher, using the findings from studies 9, 10 and 11, produced a handbook procedurally describing the design tools entitled, *Interaction Design Tools for Product User Interfaces*. This document is included in Appendix 13.5. It is recommended that the reader is familiar with its contents before reading this chapter.

Within the handbook, the purpose, function and relationship of each design tool was defined. The handbook contained five sections.

1. An *introductory section* (pages 1.2 to 1.4) briefly describing the scope of the tools, how they had been produced and a 'map' (see Figure 9.1) describing the relationship of the tools to each other.
2. *Card sorting tool* - (pages 2.1 to 2.23) This section described the function and role of this tool, how to plan and manage card sorting



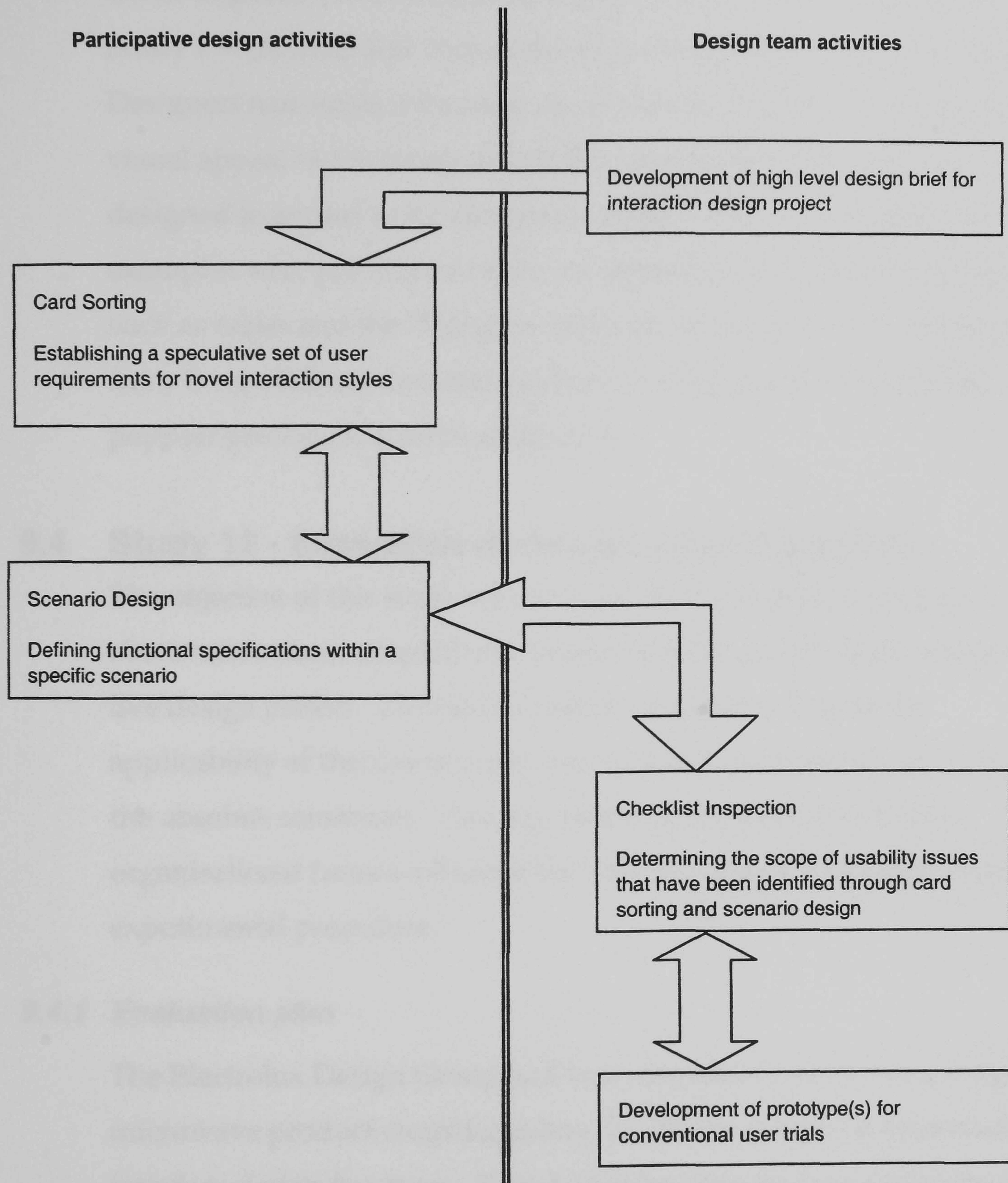
exercises, how to analyse data and how to produce a user requirements brief. Using findings from Studies 9 and 10, practical guidance was introduced on designing cards to reduce confusion and improve the validity of the design data gained from them. Guidance was also given on how to analyse the findings more effectively.

3. *Scenario design tool* - (pages 3.1 to 3.14) This section included information on how to design and manage a scenario design exercise and how to analyse design data. The use of 'tabs' to support a more tangible and open analysis of interaction styles. The use, adaptation, selection and placement of tabs to design a paper-based prototype interface.
4. *Checklist inspection tool* - (pages 4.1 to 4.28) This section included user, task and environmental issues as well as 'information processing' issues, checklists were provided to ensure designers covered all usability issues. Based on findings from Study 10, this tool was placed later in the design process when more complete interaction styles had been developed. A checklist structure was introduced to improve the usability and learnability of an inspection based tool based on findings from Study 11.
5. *Guidelines* - were provided on how to carry out interactive prototype evaluations (pages 5.1 to 5.7) and interface design guidelines (pages 6.2 to 6.60). These were produced as part of a requirement for the EPSRC research proposal but were not evaluated in this study as their level of effectiveness had already been studied.

A map illustrating the procedural relationships between the tools is provided in Figure 9.1. Each design tool was introduced by presenting



statements about resources required and what outputs would be produced from each tool.



**Figure 9.1 Procedural relationship of design tools**

To help support the use and implementation of the tools, step-by-step procedures were provided wherever possible, for example in how to



design cards (pages 2-5 to 2-8 in Appendix 13.5). The designer's role in the implementation of the tools was made as explicit as possible.

Other important lessons learnt from the development of guidelines in Study 6 were included, such as the importance of good graphic design. Designers had made comments about the importance of a strong overall visual appeal to encourage acceptance and usage. The handbook was designed to appeal to the designer's visual senses. Illustrations and examples were provided to improve accessibility. Information formats such as tables and the 'Things to think about' sections were retained from the guidelines document as these were generally found to be popular presentation styles in Study 6.

## **9.4 Study 12 - Evaluation of design tools in practice**

The objective of this study was to assess how the designers under observation used, adapted and interpreted the design tools during a live design project. To obtain a realistic understanding of the applicability of the design tools, experimental intervention was kept to the absolute minimum. This was essential to ensure that natural organisational factors influence the effectiveness of the design tools, not experimental procedure.

### **9.4.1 Evaluation plan**

The Electrolux Design Group had been commissioned to review their microwave product range including the design of new or improved interface design functions. Two designers, who had previously been involved in Study 9 (card sorting evaluation) were provided with an hour long tutorial explaining the design tools by working through the handbook. An outline timetable was drawn up with proposed dates to conduct a series of workshops. A cooking scenario was selected which the designers felt would require demanding challenges in using a microwave, such as planning when and how to use the microwave. The



meal selected was a fish curry (see page 9 in Appendix 13.6) and a crème brûlée.



**Figure 9.2 Participants involved in card sorting exercise**

Before the first workshop the designers were asked to prepare for the study by obtaining participants and preparing cards. They were asked to re-read the handbook to help them plan the cooking activities and were encouraged to consult the researcher if any aspects of the preparation process were unclear. One designer requested another meeting to assess their completed work and advise them on future directions. Suggestions on improving the design and labelling of the cards were given. To help the designers' confidence about conducting the workshops, a dummy walk through was carried out with the designers talking through the procedural steps required in each exercise with the researcher. This took about an hour.



Two card sorting workshops were carried out, each with five non-design employees of Electrolux. One workshop per day was conducted over two consecutive days employing the same designers. Figure 9.2 illustrates a card sorting workshop in action. Workshops were video recorded with the participants' consent.

During the card sorting workshops, the designers recorded any thoughts or comments on a large flip chart and photographed key events such as the completed task map. In contrast to the first card sorting exercises in Study 9, the designers facilitated the workshops and made amendments and improvements to the card sorting exercises between the first and second workshop. After the workshops, plans were set in place to implement the next design tool. This had been planned to occur after a two-week interval but was unfortunately delayed by two months due to project demands at Electrolux.

The key outcome from the card sorting workshops was a user requirements brief in the form of a large (1.5m wide and 1.0m) board (see image 9 on page 11 of Appendix 13.6). This was used to form the tab board for next design tool, scenario design, which was based on preferred function cards clustered into cells.

Tabs were based on an idea developed by the undergraduate students in Study 10 (see Section 8.7.2). Tabs were small, annotated sketches of preferred or suggested function variants produced from the card sorting exercise for participants to select in the scenario design workshops. Examples of tabs are circled in Figure 9.3, illustrating variant 'start' controls.



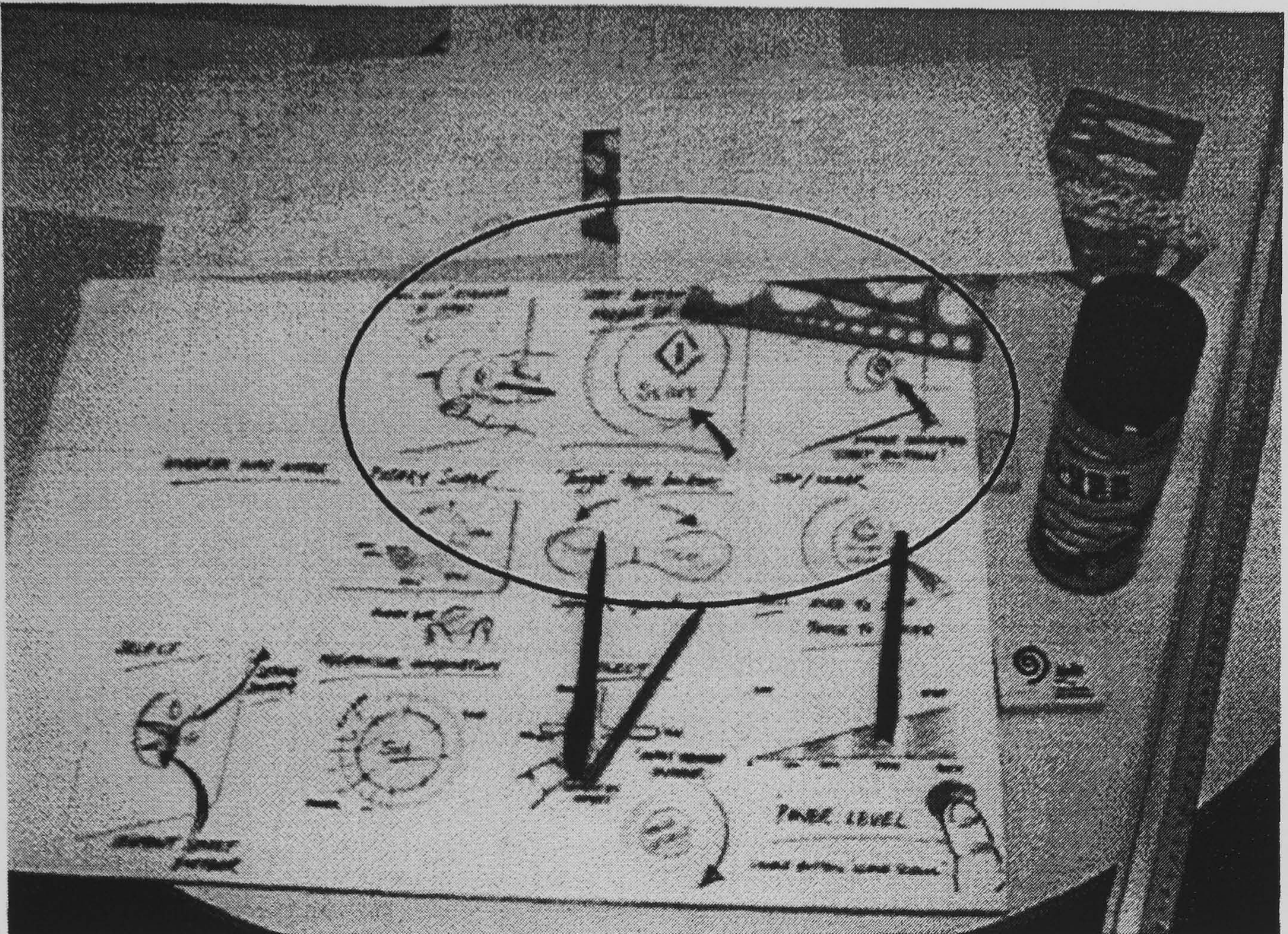


Figure 9.3 Tabs created by designers

The scenario design workshops were planned in a similar way to card sorting. The designers were asked to read the handbook and plan their workshops around the written procedures. Two workshops were conducted on consecutive days each using two volunteer participants from the previous card sorting workshops. Participants repeated the same task given in the card sorting workshop, but performed it as a real task using a working kitchen. Although they had to use a microwave oven, they could only operate the microwave 'through' the tab board and prototype interface (see Figure 9.4, tab board is at the rear of the picture). A conventional oven could not be used to force the participants to select and consider tabs (control and display components) and to discuss the usability of each component device.





**Figure 9.4** Tab board in use

Participants carried out the scenario by following a recipe and were encouraged to discuss their thoughts and ideas on the design of the proposed microwave interface. Alterations to any design proposal or tabs could be made at any time. The selected tabs were used to build up a paper prototype (see Figure 9.5) based purely on the specific needs of the participants within a given scenario. The researcher attended all workshop sessions. Video tape recordings were made and evaluated against the effectiveness criteria defined in Table 8.2.



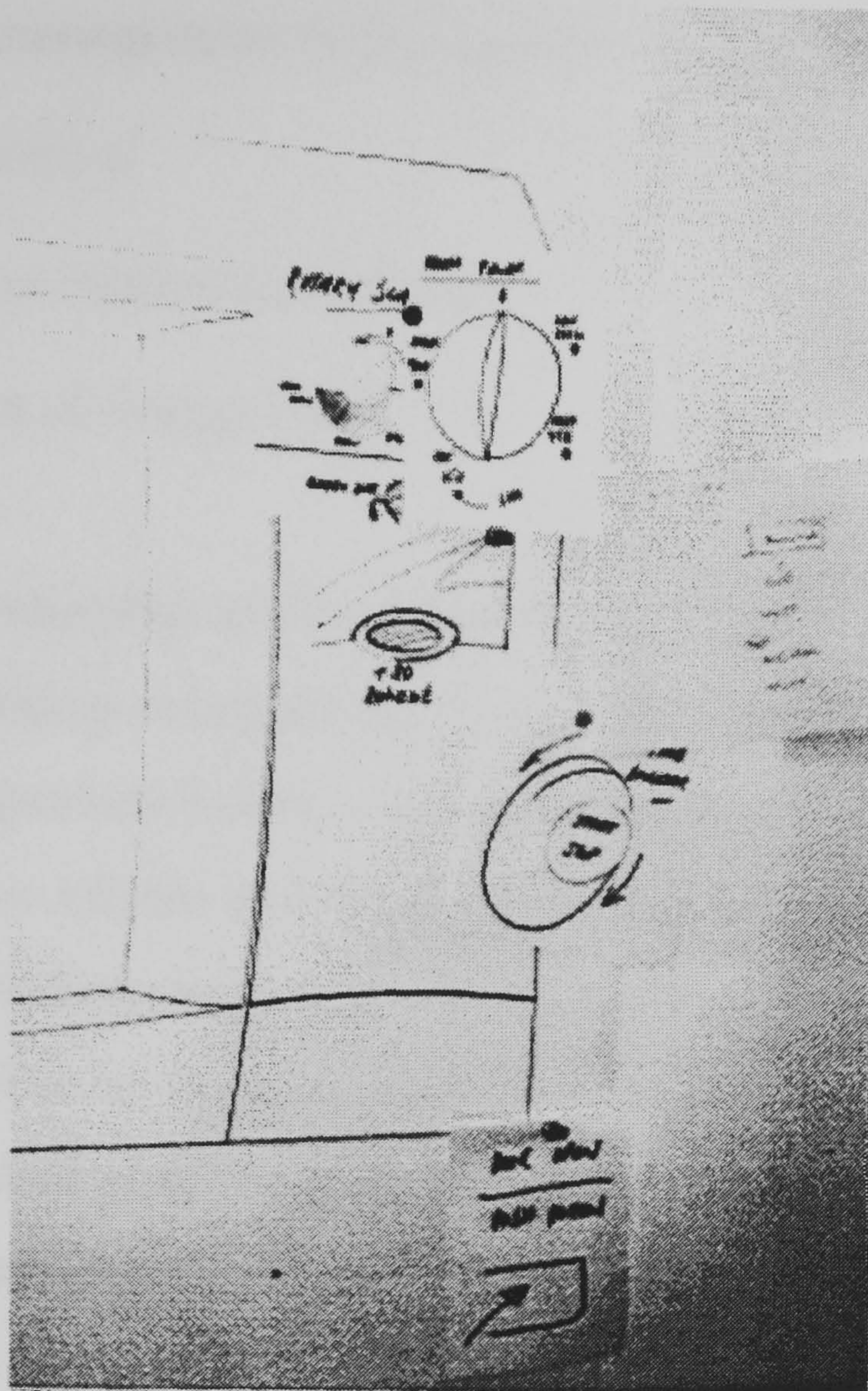


Figure 9.5 Tabs placed on prototype interface

The inspection tool was explained to the designers, but both did not feel that it would contribute to their overall aims. This tool, therefore, did not contribute to this study.

#### 9.4.2 Observations

##### *Reliability and validity of design data*

The design tools remained a very flexible, adaptable and potentially powerful device for collecting design data. However, it is precisely these important characteristics that continued to make the validity of the design data weak. Factors affecting reliability and validity of the tools were discussed in Section 8.5. Briefly the factors identified that may affect validity were:



- Design and management within design tools
- Procedural control
- Quality of experiential knowledge
- Interpretation of design data

To encourage ownership of the design tools, the designers were left to *manage* the workshop exercises for themselves. While support was provided, an important balance needed to be kept between offering too much prescriptive advice and allowing the designers to learn from experience. Providing consistent instruction to the designers was at times difficult when so many of the exercises and procedures were untested. The handbook did provide consistent instruction but was not used in preference to being given personal instruction by the researcher.

The designers did not have to *design or manage* workshop objectives. In this study the objectives were centred round research questions concerning the effectiveness of the tools. Controlling validity was not an issue with the designers, they were more concerned with ensuring the workshops appeared to be smooth running to the participants and that sufficient evidence was gathered to produce an internal report.

Although the designers were now facilitating the workshops, there was little evidence of managing collaborative design dialogue. They perceived their role as providers of design concepts for participants to test and to offer their interpretations on card depictions. Incomplete or ambiguous assumptions were not collaboratively discussed or jointly resolved. This bias was more marked in the card sorting workshops than in the scenario design workshops.



The design tools did permit different levels of design data to be *interpreted*. However, the *procedural control* within the design tools forced interaction design issues to be principally focussed at the physical device level, through a strong emphasis on depicting physical control and display elements on the cards and tabs. Other capturing methods were provided, for example, the layout of the task map provided very useful information on where participants felt that functions could support the cooking activity. This type of design data could have helped in developing navigational support. However, the designers did not use these data even after the researcher had explicitly pointed out map patterns to them. The handbook also provided explicit methods by which such types of design data could be analysed but, again, they were not used. The designers appeared to be content with physical device level solutions.

Participants were confident to criticise the *procedural control* of the workshops. In the first workshop, participants found an error in the menu instructions and sought clarification about how the cooking of two meals should be represented within the task plan.

---

D 'Why are there two lots of creme brulee?' [realising that two duplicate menus had been given to the participants]

P 'If there are two separate recipes, presumably two separate set of instructions [meaning fish curry and creme brulee], you're not instructing us to make the curry and then at some point bring the curry bit in later on... it does not mean a plate of curry mixture... but surely this should already be done?'

D 'That shouldn't be there [referring to a card] forget that..we've got creme brulee here

P 'They've got the curry over there and we've got the creme brulee over here' [laughter]

---

In this example, the participant (P) realises there are instructional errors. This leads to a discussion about how they should represent two parallel cooking activities. The designer (D) offers his *interpretation* rather than allowing the participants to share and articulate how this might be effectively achieved. Both designer and participants share in rule formulation rather than design data gathering. This can be



contrasted with a discussion in the scenario design workshop about using the microwave as a food weighing scale.

---

P1 'If I had to weigh a lot of things, I wouldn't be keen on running to the microwave all the time..'  
 P2 'because you might be in the part of the kitchen where your work surface isn't [near the microwave]  
 D 'It's something you wouldn't use?'  
 P2 'No unless you were cooking something specific by weight'  
 P1 'chicken for example'  
 P2 'but not something like this' [referring to mixed ingredients for meal they are preparing]

---

The participants (P1 and P2) explore the possibility of using the turntable as a weighing device. Decisions are based more deeply on situated task behaviour providing richer less speculative design data. By using contextual cues and experiential knowledge, participants were able to produce more situated design data to make more informed decisions about task planning and function selection. This suggested that card sorting exercises still lacked important contextual cues.

Designers created design data (cards and tabs) in quite a mechanistic way. They produced simple comparative tabs, for example different types of time controllers, for participants to comment on largely based on the outputs from the function filter used in the card sorting tool. Tab options were refined through approving, rejecting or adapting them. Producing conventional prototype solutions reduced the likelihood of more novel and user driven solutions to emerge. The designers were very effective at controlling tab selection. While participants were constructing their prototype interface, the designers often questioned their rationale for selecting or placing a tab. Often this was to highlight a syntax problem they had created but in some cases this was to allow the participants to consider alternative function variants. The designers did not record this form of design rationale. When questioned about this they felt the prototype, as an outcome, offered sufficient constructive and concrete evidence.



Many opportunities were observed where potentially useful design data emerged from *experiential knowledge* but the designers did not exploit this. There were a few reasons for this. The designers were inexperienced in collaborating with participants and were more concerned about managing the workshops than with gathering useful design data and interaction models. The designers also had little expectations about the design tools and were not seeking specific outcomes. For example, it was suggested to the designers by the researcher that cards should appear more 'rough and ready' to infer that changes and amendments could be made, thus allowing the participants to control the design of the cards. Professional pride prevented them from doing this. This had the effect of constraining the *quality of experiential knowledge* related to the meaning of the cards.

*Experiential knowledge* was obtained more effectively in the scenario design workshops. In some situations the participants evaluated the tab board (interaction design model) to great effect but again speculative decision making was found.

---

P1 'If you cook something for 10 minutes and you take it out after 9 then as long as you don't start it again there will be a minute left in your memory'  
P2 'Then you won't get your time and date back - you could leave it another day and all you would get, just sat there is 1 minute on the screen'  
P1 'So when you put your next item in to cook it will automatically update itself, the clock would run in the background'  
D 'I don't know, that's a feature you might decide to design, perhaps it clears itself after a period of time'  
P2 'Yer, it gets bored'  
P1 'auto clear'  
P2 'but that might cause problems if you go away and answer the phone'

---

In this dialogue, a new 'automatic re-setting' function was considered but decisions become based more on anticipated future behaviour rather than based on the context of the scenario. While this can be considered useful design data, it weakens the validity of these data, as it is not grounded in *experiential knowledge* from the scenario.



However, in another example grounded *experiential knowledge* was used effectively. One participant struggled with the concept of representing weight on a display and suggested that it be in the form of 'bags of sugar' as this was more meaningful as she did not instinctively know how much a kilogram would weigh.

The *interpretation of design data* was sometimes arbitrary. Participants made too many irrelevant decision making tasks that were not contributing towards effective design data. The card sorting tool placed too many cognitive demands on the participants. The selecting, placement and planning of cards inhibited useful consideration of real scenario-based activity. Decision making often related to stereotypical assumptions, for example,

---

P 'Well I could go for something as complicated as that, but I don't do a lot of cooking, but most women don't want a lot of buttons on a cooker, they want to turn it on and use it, you might be able to have a multifunction control but with a turn knob'

---

or based on personal preferences rather scenario-based demands. This was particularly true while using the function filter.

---

P1 'We'll need a START or STOP won't we or STOP AUTOMATIC?'  
 P2 'No I have to...its on, you normally click it up on mine'  
 P3 'When it gets back to zero it switches off'  
 D 'The bell rings'  
 P2 'Yes, you can stop it half way through'  
 P1 'So we are saying we don't want a manual - we'll do it through the timer'  
 P2 'So it's 'quite useful' that we don't want to use It'

---

Here, automatic features were considered but references were made to personal products rather than the scenario. The function is also inserted into the wrong bin by mistake by misinterpreting the meaning of 'quite useful'.



The graphical representation of tabs (control knobs and dials in this example) did cause problems with *interpretation* particularly temporal aspects. Often designers had to explain how a function would operate.

---

P 'Is that ... does that operate the dial?'  
 D1 'Set it and then activate it'  
 P 'Right, so you set the dial and then activate the dial'  
 D1 'No that's actually on the screen, it's above the dial in the screen... so you turn that round'  
 P 'So that's the knob on there, is that what you are saying?'  
 D1 'No'  
 D2 'Imagine that's your display it would be there'  
 P 'Oh Sorry'

---

In this example the participant has confused a display having an illustrative icon on it with a real control object that could be physically manipulated.

Reliability of the design tools was also examined. Again, the factors discussed in Section 8.5 were used, these briefly were:

- Consistency of design tool management across workshops
- Changes in procedural understanding
- Consistency of experiential knowledge

Small participant groups affected the consistency by which the design tools could be *managed*. The small number of workshops also prevented measuring the *consistency of experiential knowledge* gained to any great extent. One of the observations made by the designers was how workshop outcomes differed between group members. It was suggested that this could be overcome by having an intensive 2-week period where workshops were more consciously manipulated in terms of participant composition and activity selection. This was an issue that Danfoss in Study 8 were very keen on. Nevertheless, similar participant behaviour between workshops groups was observed. They



sought clarification for rules and interpretation, particularly during the early stages. In the card sorting exercises, participants used only the cards made available to them, thus fitting their interaction model around the designers' concepts.

Adaptability of the design tools has been identified as an important aspect of *managing* the tools and was encouraged. Providing this form of control did, however, have an effect on the *procedural understanding* of methods adopted thus affecting the reliability of the outcomes. Between the first and second card sorting exercise, the designers recognised by themselves that the first task planning activity was too detailed and procedural. The designers had produced cards with prescriptive instructions resembling instructions from a cooking recipe. They also recognised that card depiction was 'text' heavy and more graphical images would increase interpretation and improve card recognition. However, many of the amendments that were introduced by the designers did not have a major impact on reliability. Most of the changes made to the procedures were implemented to reduce decision making the participants had to make rather than affecting the type of decisions being made across different workshops.

In considering the importance of these criteria, they need to be set against the real objective of using these design tools. Reliability and validity only become important if consistent and robust interaction design issues are required. If, on the other hand, the objective is to generate creative interaction design solutions for further discussion or refinement, then the rigour of these methods is less important. It was not possible at the outset of this study to know what the design tools would be most suited to. Certainly the designers were content with the tools providing them with creative design data.



*Interpretation of interaction design models*

In interviewing the designers after the card sorting workshops, they appeared to have a very clear view about what the participants wanted from a microwave interface. They suggested that the participants were reluctant to use non-tactile control devices, as one designer stated, 'that means no more touch screen interfaces'. The designers concluded that any proposed interface should have no more than three control devices, although this had not been discussed during the workshops with the participants.



**Figure 9.6 Example of task map and function filter**

The function filtering activity was thought to be the most productive interaction model in terms of establishing design guidance in the form of user requirements. Although the task map and function filter were photographed (see Figure 9.6), they were not analysed to produce further types of design data such as gaining a navigation model from the task map (bottom left hand corner of Figure 9.6). Only the preferred



cards from the function filter (top right hand corner of Figure 9.6) were retained to form tabs for the scenario design workshops. The effectiveness of the function filter was therefore strongly dependent upon the representation of functions on the cards and the level of critical judgement that participants applied to evaluating function. The designers used the preferred function cards to form user requirements and these were presented in the report using the tabular structure suggested in the handbook. The table describes each function in terms of their importance, frequency of use and their significant of use between different types of users and tasks (see pages 18-30 in Appendix 13.6).

With the next interaction model, the tab board, both designers and participants were more critical in their evaluation, selection and use of the tabs. New tabs were devised if the participants identified an alternative or improved way of achieving a goal. In some situations participants made requests for functions that did not exist on the board. In this situation, the designer and participant would simply draw a new tab that suited their needs. The process of design, build and test could be achieved in a matter of minutes.

Cards and tabs often contained too much design detail. For example, a card might have a function such as 'Determine Cooking Time'. This would be associated with a graphical image of a clock, followed by instructional information 'Enter time using analogue clock face'. Many detailed aspects of an interaction style were therefore embedded in card depiction. Most of the control and display tabs offered by the designers were combined features or functions. This restricted the options available for the participants to select specific control and display components. Cards with this detail reduced creative discussion. Nevertheless, discussions about personal cooking habits and preferences occurred on an ad-hoc basis but had little impact on the



interaction model. This problem was realised by the designers and was improved for the following scenario design workshop by using more interpretable tabs.

### *Scope of usability issues identified*

The intention was the designers would review the final design solutions using the inspection design tool to ensure that such issues were addressed. However, both designers were reluctant to use the tool explaining that such investigative detail was unnecessary, as only conceptual design proposals were required at this stage. The use of the inspection design tool was therefore abandoned for this study.

As the inspection tools was not used the ergonomics criteria (as defined by Scapin and Bastien 1997, see Section 2.5.3) were used to identify the scope of usability issues that were identified. Only issues related to task and user compatibility were raised with the card sorting tool.

Usability issues were more comprehensively addressed with the scenario design tool, particularly issues such as consistency, the significance of codes and compatibility. Participants questioned control labels and provided more meaningful labelling suggestions; for example replacing power levels, described in Watts, to a more concrete coding method using values such as full and half power. There was also evidence of error management assessment, particularly with syntax errors through the extensive use of walk through testing in the scenario design workshops. Error management was introduced by the designers at the card sorting workshops by introducing 'event cards', in one situation participants were asked to re-consider their task map by having a new card inserted labelled 'Not defrosted properly'.

Participants were asked to review their function cards and check they had adequate functionality in the map to manage the new event. In the first workshop this triggered a review of their function cards while in the second the outcome was less resolved.



---

D 'You find the fish has not defrosted properly [inserts card] what would you do then? What function would you use and how long would it now take to defrost?'

P1 'Just have to guess then'

P2 'That's why we wanted the 'weigh' one [referring to a function card they had considered earlier] so we would not have to keep checking it when we took it out'

---

The interaction model was not sufficiently robust to handle detailed interrogation. Both design tools were poor at identifying issues such as guidance, workload and adaptability where dynamic and temporal aspects of interaction were difficult to model.

### *Supporting novel interaction styles*

A range of novel concepts was introduced by the designers with the card sorting tool such as twin turntables, ready meal scanner, oven management system, universal input controls and menu cards. Most novel concepts were quickly rejected at the card sorting stage. Some further degree of novelty in interaction styles was introduced in the scenario design, either by participants making requests for functionality that was not provided on the tab board, or by collaborative discussion with designers. Novel suggestions at this stage, however, were through the adaptation of control and display features already provided as tabs rather than the consideration of radically new interaction styles.

The design solutions produced by the design group from this study were presented in an internal management report. This report was submitted to senior design management for approval and to contribute towards a product development strategy. The report did not exhibit many of the radical novel interaction design concepts, certainly none that were gained from using the design tools. The final proposals retained many concepts and ideas that the designers had been discussing before the introduction of the design tools. However some user requirements from the design tools were implemented. For example, labelling the power control using Watts was retained, despite



the participants' dislike for such a feature but mechanical, physical controls were retained, a strong preference in the card sorting workshops. The report (Appendix 13.6 pages 35-42) offered five interface proposals including a 'lifestyle snack product, mid and top range functional specification interfaces and an interface offering remote control.

### *Usability of design tools*

Despite ample opportunity for the designers to read the design tool handbook and to become familiar with it, many instructions or suggestions within the handbook were ignored. General procedural principles were adopted but many of the subtleties, particularly in card and tab design were not implemented. Nearly all guidance and learning of the design tools came through instruction and experiential learning rather than reading the handbook. Both designers were hesitant and unsure about user participative methods and were keen for external advice in the development of the workshops as much as possible. However, during the workshops, the designers did use their own initiative without seeking advice from the researcher assisting in the process and spontaneously resolved many procedural problems as they occurred.

The designers quickly gained confidence in using the tools and adapted the tools to improve their performance with them. Once the first workshop was complete, both designers expressed a clearer understanding of managing and conducting workshops. They identified improvements and made a series of suggestions, for example replacing the matrix filter. The handbook was very poor at providing accessible or understandable ways in which designers could record design data or analyse interaction design models and this needed to be improved. The only effective currency in terms of design data was the physical device-based functions depicted either on cards or tabs.



Evidence for this observation can be seen in the design report where references are only made to the card based functions. The selected 'preferred' functions from card sorting form the basis of a rationale and direction for the interface design proposals.

### *Relevance of design tools to designers*

Interviews with the designers and observations provided evidence the designers had taken ownership of the design tool. Although the first workshop proved an exhausting process; as one designer said after the participants left, 'that was absolutely knacker!' They enjoyed the experience and thought it worthwhile as they had gleaned design data they had not expected, particularly on preferences towards traditional controls that could be embedded with advanced interactive technology.

Initially, both designers expressed concern about the involvement of participants in design decision making. They were wary of suggesting vague design ideas to participants in fear of appearing unprofessional or inexperienced. This concern quickly evaporated once participants and designers became familiar with their roles.

Although the relationship between the designers and participants was more consultative than participative, certainly in the card sorting workshops. External support and advice throughout was very important to the success and understanding of the design tools but it was also very evident that the designers progressively gained in confidence and enjoyed using the tools. One of them was considering using them on other non-interface design projects. The designers were positive about the outcomes and thought they had gained useful insights that would not have been otherwise gathered. They considered the handbook to be a useful resource but would have preferred some form of quick reference guide in the initial stages to help manage workshops.



*Likelihood of organisational survival*

The design manager and senior designer in the group were interviewed to discover how they perceived the efficacy of the design tools, the quality of the design data and interaction models and the quality of the final design solutions through reading the management report.

The design manager was extremely encouraged by the adoption of the design tools and felt they reflected a recently implemented product development philosophy known as Integrated Product Development. This process was driven by the organisation's 'core values' including a user-centred approach to product development rather than on historical organisational production methods. The introduction of the design tools was also regarded as timely as the role and skills of the design team was beginning to change by developing more innovative and user-led product proposals within the Electrolux group. In order that the design tools could gain greater acceptance, approval needed to be sought at a senior management level. This was the key purpose of the internal management report (Appendix 13.6). He suggested that the design tools would only survive if the final design solutions were sufficiently creative and in line with current product development requirements. If this could be proven the design tools needed to be 'sold' outside of the immediate design group in 'power point' presentation terms. It was important the design tools could be explained and implemented succinctly to other disparate and culturally different design groups within the organisation. This, he thought, could only be done through effective training and not through text-based documents which should only be provided as reference material. He added that the design tools would have to compete against any number of 'tools' used within the IPD process.

The senior designer expressed the importance of adaptability and again referred to the organisational changes that were currently underway.



His view was that design skills were going to be dissipated through the organisation and designers would become part of 'development groups'. In this sense the design tools would have to be acceptable to a much broader skills base. He also agreed that the design tools could only survive if they could be integrated within the IPD process. There was little incentive for any member to use methods not recognised as part of this process. Tangible and cost-effective benefits for introducing design tools needed to be evident.

### **9.4.3 Reflections**

The use of the design tools was clearly enjoyed by both participants and designers and acceptance for their use had been achieved. The designers were very encouraged by outcomes from the design and many of the participants' proposals were translated into the design management report. Unfortunately there was little evidence that the designers had been critical or objective about the effectiveness of the tools. They enjoyed the design process but felt little compulsion to relate their final design proposals back to the outputs from the design tools. Furthermore, at a management level, further evidence was needed of their creative power and marketability before they could be accepted more broadly in the organisation.

Of all the effectiveness criteria reviewed, assessing validity caused the most concern. Allowing the designers to manage and control the design tools meant there was less direct control over outcomes from the design tools. Experimental control moved from the design and adaptation of the design tools to a form of remote monitoring and instructing. Through instruction, the researcher attempted to redeem some of the deficiencies that had been identified in earlier studies. Designers, for example, were encouraged to use more ambiguous, less detailed cards to prompt more open and potentially more fruitful design data. This advice was largely ignored in favour of cleanly composed cards and



tabs for reasons that were due mainly to professional design pride. This resulted in some of the less defined and potentially more interesting concepts, such as cooking management systems, having less currency or weight in the workshops. The designers did recognise this and as they became more comfortable with the design tools, were prepared to use more 'rough and ready' material.

Although validity of design data was a concern, the interviews with the design managers suggested that this would be of less concern to them. Effectiveness of the tools was viewed entirely on the quality of the creative and commercial solutions the design tools would provide with little regard given to how this might have been achieved. Important issues were also raised about how these design tools might be organisationally implemented. Training in the use of the design tools was clearly seen as an obvious route. The tools had to satisfy complex organisational demands, be marketable, quick to learn but also produce commercially appropriate design solutions. More consideration needs to be given to how this could be accomplished.

The introduction of these design tools satisfied the applicability criteria of *usability* and *relevance to target audience* but would not, in their current state satisfy the criterion of *organisational survival*. The specificity criteria of *reliability and validity* still needed to be improved but other evidence suggested that these might not be of high importance to the designers' overall objectives. The tools needed to have more implicit procedural controls to ensure that they addressed usability issues in a valid and thorough way. To explore how these particular criteria could be improved, the design tools were presented to another organisation.

Shortly after this study, three of the designers used in this research left Electrolux over a period of about 6 months. In informal discussions one of the designers suggested that they had become frustrated by the lack



of influence they were having individually and as a group in product development. One of the designers had used examples of this research work in a subsequent job interview.

## 9.5 Study 13 - Investigating applicability factors further

An opportunity arose where the effectiveness criteria *organisational survival* and *reliability and validity* could be assessed within another product development organisation. A discussion group session was set up with members of the User Centred Design Group at British Telecom Research Laboratories, Martlesham where the feasibility of introducing the design tools into future user-centred design projects was considered. The user-centred design group was composed mainly of industrial designers.

### 9.5.1 Evaluation plan

Members of the User Centred Design Group were invited to a one and a half-day workshop. Four members of the group plus a member of the human factors group volunteered for the workshop sessions. These included three industrial designers, an ergonomist, and a placement, post-graduate student studying Human Computer Interaction. The objective of the workshop was to identify how *reliability and validity* and *organisational survival* could be improved.

An introductory presentation was given to provide an overview of the design tools. The presentation included background information on the outcomes of the Electrolux study, along with identified strengths and weaknesses. Rather than describe the tools as a rigid procedural and prescriptive format, alternative suggestions on how the design tools might be adapted were presented for discussion. Focus was placed on how best design data and interaction models could effectively be generated within their organisational context. Specific emphasis was placed on the group critically assessing their perceived effectiveness in



gaining outcomes relevant to their organisation and in terms of fitting them into their design and organisational culture.

As the workshop was primarily a small investigative study, producing tangible design proposals was not a feasible objective. It was not possible to consider the scenario design tool in detail, as potential users or participants were not involved in this workshop. Also, the workshop was not related to a specific project, participants were therefore asked to consider a current or future project, which could be used as a case study or reference point for discussion throughout the workshop.

### **9.5.2 Observations**

After the introductory presentation, a design case study was discussed and agreed upon. It was decided to use a recently conceived conceptual project one of the designers was currently engaged in. The aim of the project was to investigate possible future functions and interaction design issues for an existing hand-held testing device used by BT field engineers. This device was used to assist in installing and testing new telephones. One of the possible future functions of the device was to support the engineer in undertaking sales activities.

The group were then asked to draw up a list of critical success factors that they thought were important for the design tools to be effective in the development of this new device for field engineers. The group decided that any proposed design tools should provide a good overview of the engineer's job, the engineer's perceptions of the job and they should find using the tools acceptable and meaningful. This was thought important to facilitate an increased involvement of end-users in the design process. As far as the design group was concerned, the tools must be acceptable to many different stakeholders. The tools must allow existing organisational human factors and user-centred skills, methods and knowledge within BT to be integrated while also



permitting other non-design disciplines to be involved in the process. The use of the design tools should be seen as viable and acceptable to senior management and particularly marketing groups. Finally, the outcomes of the design tools must translate naturally into tangible design specifications.

Once the card sorting tool had been explained, the designers decided to go through an exercise to familiarise themselves with the procedures. The group decided to generate cards while assuming the role of a BT engineer. One of the key issues that needed to be explored was how the engineers would accept the role of sales advisor while installing BT equipment. Therefore, the card sorting exercises needed to focus on events where the engineer was in contact with the customer. The group suggested using graphical images, which emphasised communication between the engineer and the customer. A set of cards was produced depicting images such as a book for documentation and a customer door to suggest making contact at the front door. Then, while playing the role of engineers, the group began to write down statements or brief task descriptions on the cards, prompted by the graphical images. Once the statement was written, it was then called out by each participant and thrown into a pile in the centre of the table.

When a reasonable number of cards had been produced, the group then sorted the cards into a map so that events or activities presented a story or temporal record of a typical site engineer visit. The cards were placed on the table from left to right. Once this was complete the group then commented on the map to establish possible design guidelines. One of the group members suggested that cards could be sorted using coloured circular stickers to aid in classifying common card themes, for example, critical events.



The group was then asked how this map could be translated into an interaction design model that they felt, in their role as designers, would allow them to consolidate design data. User requirements were discussed and some were suggested, for example, 'the device must not stop me from getting on with the job quickly'. Other suggestions were put forward such as, importance/frequency tables and 'likes'/'dislikes' bins similar to those used at Electrolux. These were dismissed, however, in favour of bins labelled with 'trigger' statements to prompt participants to articulate useful contextual design data. Some suggestions for triggers were presented such as 'Things that make my job difficult', 'Typical conversations with customers', and 'Job satisfaction'.

The group was then asked to consider how problems of abstracting design data could be dealt with. Design data could, in effect, be raised at a number of levels, from preferences towards physical control and display devices to different procedures in carrying out a task to cultural and organisational factors affecting job performance. The group was asked to consider how information generated at different levels could be captured, controlled and represented. Two key suggestions were made. First, simply by posing questions targeting different levels of interaction design data. Second, users and designers collaboratively obtain design guidelines covering as much design data from the card sorting exercises as possible and then reviewed by the designers for identified gaps.

The discussion session concluded by summarising how the design tools could be improved or what type of new design tools could be developed (see Table 9.1). It was suggested that design tools should focus on retaining links between inquiry objectives (what type of design data is being elicited) and outcomes (how interaction design issues could be best support these objectives). Design tools, and exercises



within the design tools, were suggested that might effectively generate relevant design data and analytical questions were suggested that might support appropriate design data abstraction.

During the discussion, other hurdles were identified that might impede survival of the design tools. Most of the designers worked closely with human factors specialists and traditionally user requirements capture was regarded more as a human factors role. The designers needed to feel confident about using alternative design methods, which they perceived as having a human factors philosophy, before exposing their design methods to external criticism. Some concern was expressed about 'treading on human factors territory' and some of the group felt unsure about being able to provide a robust design rationale for any design proposals using this approach. Concern was also expressed about being able to gain access to users for workshop sessions. Designers were usually co-opted onto observation sessions and focus group meetings organised by other sections of the organisation. The group rarely organised participative sessions themselves. The BT group's poor access to participants or potential end users was an unanticipated problem. User-centred design methods were more commonly found in the human factors group, not the User-Centred Design Group!



Objectives	Possible new or adapted design tools	Possible questions posed to obtain valid and reliable design data	Examples of anticipated outcomes - suggested by the group
Find out what field engineers do	Field Observation  Contextual Inquiry	What are the tasks, procedures, terminology and artefacts used?	None
Find out how field engineers perceive and understand their job	Development and comparison of task maps using different workshop groups  Use cartoon cards and sort into bins with 'trigger' statements  Categorising or themed cards	What aspects of the job need to be preserved?  What are the critical parts of the job?  What concerns do the engineers have for future changes to their job?	I would like to have control over the order in which I do my calls
Find out what type of equipment they use, peripheral artefacts e.g. personal note books What can be changed?	Artefacts cards in bins such as things: - I really need - Company provides - I provide - I never use	What works and what doesn't? Why do they like using them or not? What could the new device replace or enhance How might the new device alter their job	The 'job' notebook is very important and used every day. This would need to be part of the new device.
Find out how they currently communicate with customers.  What are their experiences and attitudes towards customer contact?	Use prompt cards that promote stories or anecdotal evidence	What do they normally talk about? What do the engineer and customers value?	Customers prefer me to get on with the job
Find out who they communicate with, how and frequency	Use prompt cards for different scenarios such as:  Communicate with: How?: By: How often?:  Then build a spider diagram	What communication links are critical?  What about social communication?  What type of impact would the device have on communication flow	It would be good if I could have a 'hot' link to an expert on 'inserting telephone plugs'
Find out what type of technology would be acceptable	Ask participants to insert concept functionality cards into completed task map using a mixture of sketches and statements	How much should the device support sales activities?  What types of interaction styles are appropriate?	A 'sales' device could be given to the customer so the engineer can be carrying out the job while the customer completes and interactive questionnaire

Table 9.1 Summary of suggestions to improve design tools

9.5.3 Reflections

Overall, whilst the designers demonstrated a keenness and willingness to use some of the exercises within the design tools, demonstrating the



benefits of the design tools in a contrived fashion without a clear design objective and without the involvement of end-users made the complete acceptance of the design tools difficult. Some of the participant members had come to the workshop with different expectations. Some members of the group had anticipated gaining instruction in fully developed and proven design tools. However, the group provided useful suggestions to how these potential design tools may survive in BT.

Similar concerns were expressed to those in Electrolux, effectiveness of the design tools was not judged on the validity of the method but on how effectively the outcomes could be disseminated across the organisation.

## 9.6 Conclusions

In the BT study, an interaction design problem was contrived to create a point of focus for discussion. This resulted in less obvious objectives and outcomes, which prompted a more critical debate about developing effective design tools. This was in marked contrast to the Electrolux study where debate was aimed at procedural matters such as improving card design. The use of a live design problem may have defined the problem space more but nevertheless, the BT debate provided evidence that design tools needed to be explained through case studies and examples as well as through walkthrough demonstrations.

Observations from both organisations indicated that both design groups were continually competing against other functionally related groups. In this competitive environment, new design tools have to quickly suggest their capability in terms of problem solving, usability, adaptability, robustness (against scrutiny from other sectors of the host organisation), and provide organisationally relevant outcomes. The BT



workshop built on some emerging thoughts from the Electrolux study. Design tools need to provide support beyond their immediate functional purposes and support other auxiliary needs. For example, outcomes from design tools must map easily to organisational reporting structures such as project reports or even 'Powerpoint' presentations to improve their viability and competitiveness against internal rival sections of the host organisation.



# 10 Conclusions

---

## 10.1 Introduction

In this final chapter, the research findings from this inquiry are reviewed and then placed in the context of other recent research. The findings from this inquiry are then critically discussed and their contribution to the field of interaction design knowledge highlighted. The chapter concludes with an outline suggestion of future research for the further development of effective secondary design tools.

## 10.2 Inquiry review

The thirteen studies in this research form a progressive and formative account of how to improve the effectiveness of interaction design tools. The early studies attempted to improve the effectiveness of design guidelines by focussing on applicability by developing them through user-centred design methods and improve specificity by proactively developing guidelines with specific relevance to CPT.

It is in the accumulative evidence gathered throughout all these studies, combined with supporting evidence from other related fields such as HCI and ergonomics that strongly indicate that guidelines in any form have marginal impact on interaction design activity. Introducing designers to tools providing attenuated interaction design data has been demonstrated to alter design behaviour. This research has demonstrated that design data has to be produced within a situated context and cannot be produced externally by the research domain. Knowledge and useful design data created by a design domain cannot conclude their published findings in the form of design guidelines. A further step is required to ensure that these design data become



implicitly integrated into the practice domain through the use of secondary design tools.

This requires in-the-field research activity of the type used in the latter studies in this research. Prototype design tools (incorporating design data from the research domain) need to be tested within a practice-based context using interaction designers on commercially active design work. Design environments are so complex that any other form of analytical study will simply not meet all the effectiveness criteria developed in this research. Initially, design tools are likely to be bespoke to a particular organisation but over time general characteristics or re-usable components of the tools maybe identified.

Design researchers will need to collate and synthesis these state-of-the-art methods and techniques devised by related research disciplines. Theoretical and applied methods would be reviewed and provide a creative springboard from which new adapted design tools could be produced.

The next section provides evidence of the use and dissemination of design data is still very traditional.

### 10.3 Recent research

Recent published papers (1999 - 2001) in the following journals were reviewed: *Communications of the ACM*; *Creativity and Innovation Management*; *The Design Journal*; *Design Issues*; *Design Management Journal*; *Design Studies*; *Ergonomics*; *Human Factors*; *Human-Computer Interaction*; *Interacting with Computers*; *Interactions: New Vision of Human-Computer Interaction*; *International Journal of Human-Computer Studies*; and *Personal and Ubiquitous Computing*. The objective of this short literature review was to identify and critique recent research, which has *direct* relevance to the findings from this research. Other studies, identified as



potentially having a contributing role, were also highlighted in consideration of future work outlined in Section 10.5.

### **10.3.1 The use of guidelines and standards**

In the early stages of this research, one of the reasons for pursuing the development of effective guidelines and standards for CPT was due, in part, to their prevalence and importance in HCI.

This is still the case. Guidelines continue to be published, see, for example, Bevan (2001) and Maguire (1999). Also, HCI standards remain little more than design principles. Bevan (2001, p533) states, 'Very few of the HCI standards specify the interface precisely, instead defining general principles from which appropriate interfaces and procedures can be derived'. Research addressing the improvement of guidelines still places emphasis on improving *specificity* through different approaches - for example, by identifying appropriate sources of guidelines (Carter 1999); improving tools to collect, manipulate and use guidelines (Vanderdonckt 1999); or by making them more context-specific (Henninger 2000) by using case-based and evolving organisational knowledge. *Applicability* factors have not been a central concern within these studies.

Alternative approaches to the design management of interface usability have been considered through the recent introduction of ISO standards for user-centred design using design management and organisationally oriented approaches (ISO 13407:1999 *Human-centred design processes for interactive systems* and the associated ISO TR 18529:2000 *Human-centred lifecycle process descriptions*). Their effectiveness, however, in terms of applicability remain in doubt. Earthy, Sherwood Jones and Bevan (2001, p 569) state that, in terms of these standards being successfully implemented, '... research is required into barriers of the uptake of user-centred approaches within technically driven engineering



disciplines'. The studies in the latter stages of this research have begun to establish where those barriers of uptake may occur. (See section 9.5)

### **10.3.2 Use of design tools**

Two studies attempting to improve the implementation of design tools and having specific relevance to this research were found - Andre, Harston, Belz, and McCreary (2001) and Säde (2001). A critique of their studies was made, using the effectiveness criteria and research findings from this inquiry.

Säde (2001) provided a descriptive account of introducing user-centred design methods into a design consultancy. He initially took a broad range of methods and techniques from the HCI field and, through iterative refinement, adapted and evolved a sub-set of these techniques to form a user-centred design toolkit for industrial designers. He made similar observations to those resulting from this research and found a lack of 'systematic' investigation in introducing HCI methods into design organisations. Säde found that design tools needed to have assistive or supporting procedures, for example to help the designers to communicate with software developers or to help participants to remain focussed during user-centred design activity and that collecting data from live design projects was problematic due to commercial demands.

Säde also introduced adapted HCI techniques such as cognitive walkthrough, paper prototyping and guidelines. As with this research, early design tools began as literal translations of HCI methods but quickly evolved as the designers began to adapt them to make them 'streamlined, to be more informal, conversational and lightweight' (p 27). To ensure that all usability issues were considered, he introduced a usability scoping tool, which evolved using design guidelines based on heuristic evaluation methods (Nielson 1993). The final toolkit had



seven techniques: contextual inquiry, affinity diagram, bridge for buttons, paper prototypes, focus groups, usability testing, guideline walkthrough, and software prototypes. These tools spanned more stages of product development than the tools developed in this research and ranged from initial data gathering through observation and interview (contextual inquiry) to summative usability testing. The design tools in this thesis had been developed on the assumption that exposure and involvement with end-users was very limited (Study 7).

Specificity Criteria	Applicability Criteria
<i>Reliability and validity of design data</i> - Not formally evaluated, context implicitly recognised	<i>Usability of design tool</i> - involving designers in the iterative development of the tools meant they were 'streamlined' to the designers' needs
<i>Interpretation of interaction model</i> - Used affinity diagram, recognised as an important mediating device between designers and software developers	<i>Relevance to proposed target audience</i> - acceptance achieved but not by all designers
<i>Scope of usability issues identified</i> - Dealt with by a set of design guidelines based on heuristic evaluation	<i>Likelihood of survival or usage in organisational context</i> - adoption of tools resulted in recruitment of designers with stronger interaction design skills
<i>Supporting novel interaction styles</i> - Used on design projects with novel interaction styles	

Table 10.1 Critique of Såde (2001) study using effectiveness criteria

In Table 10.1, critical comments are presented against the effectiveness criteria. Effectiveness was primarily measured by the relevance of the design tools to the proposed target audience; 'taking fast, lightweight, even sketchy usability action is more reasonable and beneficial than either conducting in-depth research or doing nothing at all' (p 30). Although Såde recognised streamlining might reduce the validity of the design tools, the potential impact of this was not discussed. The only



defence for the design tools was they made the designers more employable for designing smart products.

In the study by Andre *et al* (2001), the reliability of design data in the software development industry was systematically assessed. Even in the relatively mature software development practice, the authors contend that the reporting of usability information was inconsistent and needed to be improved. They state that usability information was 'conveyed by memory and word of mouth to designers responsible for fixing the problems' (p108). This resulted in poor returns on investment in usability testing and therefore the reporting and recording of usability problems requires improvement. The authors proposed a 'user action framework', which unifies and integrates a series of usability tools to support different stages of the interaction design life cycle.

The authors went on to develop a set of tools. The Usability Design Guide tool supported early stages of interaction design by providing guideline-based help. In the next design stage, usability testing, the Usability Problem Inspector and Classifier tool was used during which usability problems were identified and then classified and retained within a database. In evaluating the Usability Problem Inspector and Classifier tools, one of the hurdles to measuring reliability was found to be the degree of variation amongst the way in which subjects classified usability problems. They refined this tool to accommodate variation. Reliability had to measure consistency in classifying outcomes rather than the decision making processes that led to the outcomes. The evaluation of reliability in this study was limited to the classification of known usability problems using common problem descriptions.



Specificity Criteria	Applicability Criteria
<i>Content emphasis</i> - based on Norman's (1986) theory of action model, context of interaction not emphasised	<i>Usability of guidelines</i> - Not explicitly evaluated
<i>Structure and presentation</i> - Web based database, data were presented differently depending on if it was being viewed to inspect or classify a usability problem	<i>Reliability and validity of guidelines</i> - Reliability extensively tested against earlier versions and heuristic evaluation
<i>Scope of usability issues identified</i> - Improves as content is added to database	<i>Relevance to proposed target audience</i> - Not evaluated
<i>Support for novel interaction styles</i> - Tested against voice-based interfaces and suggested that only minor changes to labelling was required	<i>Likelihood of survival or usage in organisational or design context</i> - Not evaluated

Table 10.2 Critique of Andre *et al* (2001) study using effectiveness criteria

The authors carried out a comparative evaluation study using usability practitioners to measure the level of reliability of classifying usability problems between their web-based decision support tool and Nielsen’s (1994) heuristic evaluation method. Improved reliability was found in their decision support tool but perhaps the more interesting aspect of this study was how narrowly factors concerning applicability were tested. Participants were given 15 controlled case study descriptions with known usability problems. As in many of the studies included in the literature review in Chapter 2, effectiveness criteria were assessed against controlled usability problems divorced from real-world design problems. Furthermore, reliability was the only applicability criterion discussed.

Table 10.2 summarises the Andre *et al* study against the effectiveness criteria used in this thesis. Reliability was treated as an external effectiveness criterion as the usability information (design data) resided as a remote resource in the form of a database. Indeed, it could be argued that reliability was evaluated within a laboratory-controlled



study and it may be more unpredictable if the classification tool was tested against design problems posed by the practitioner. Much of the content emphasis of the database was grounded in Norman's theory of action model (1986). This narrowed the content scope, as the model does not naturally consider contextually rich data. The authors did argue that contextually sensitive data could be captured within different categories of the User Action Framework, but they did not expand on how this would occur.

Comparing the two studies revealed contrasting emphases in measuring effectiveness. The Säde study emphasised applicability, ensuring user-centred design methods conform to all applicability criteria but neglecting to assess reliability and validity. In contrast, the Andre *et al* study was based on a central assumption that design data can be reliably recorded and used as an external resource. This is in contrast to the findings of this research, which suggests that externally retained design data does not have any significant impact on design behaviour. Both the Säde and Andre *et al* studies serve to underline the importance of concurrently reviewing all effectiveness criteria.

### **10.3.3 Bridging the gap using context**

In Chapter 1 the issue was raised of a gap existing between practice and research. This thesis suggests that this gap can be partly 'bridged' by researchers and practitioners sharing a 'theoretical' understanding of interaction in context using models such as activity theory, distributed cognition and situated action theory. These models remain viable descriptors (Fields and Wright 2000) and continue to have an influential role. Activity theory has been used as a conceptual framework on which to base design decisions (Macaulay *et al* 2000) and Turner and Turner (2001) have provided steps towards a systematic use of activity theory to identify design solutions. Scenario design also remains an important contextually based design and evaluation tool (Carroll 2000)



which can facilitate creative collaboration between designers and users (Bødker 2000). The use of games and role playing continues to evolve for example, the use of 'movie making' in eliciting user requirements (Pedersen and Buur, 2000).

The importance of contextual analysis has resulted in ethnographic studies being adapted for use in HCI studies. For example, the use of trained design ethnographers using a structured framework (Macaulay *et al* 2000) and the introduction of cognitive ethnography (Ball and Ormerod 2000), where certain methodological characteristics are deliberately restricted to affect factors such as objectivity and verification of data. Ethnographic studies have also been used to study real-world design practice (Button 2000) and consumer product usage (Wilenmann 2001). Despite attempts to improve the quality of design data gained from ethnographic studies, there remains a debate as to how such data can be used in the design process (Viller and Sommerville 1999).

Assumptions still remain that design data can be externally accumulated by researchers and subsequently adapted for use in practice. Sutcliffe and Carroll (1999 p 216) state, (while commenting on 'reusable claims' - contextually sensitive design statements that suggest potential ways in which usability problems can be resolved) 'The users of claims are intended to be software engineers, so another motivation for this research is to spread HCI knowledge beyond the community of human factors specialists; however, effective delivery of design knowledge is a research topic in its own right...'. In this thesis, this assumption has been questioned. The 'delivery of design knowledge' must be an implicit part of producing the design data. Researchers should play an important facilitating role in allowing this form of design data to emerge effectively.



This review of recent research confirms that many of the directions taken in this thesis still reflect key trends and concerns in HCI and have relevance in the field of interaction design. It has further shown that other research studies in this area do not take a broad enough view of assessing the effectiveness of implementing research design data into practice.

## 10.4 Critique of thesis

Methodological issues are addressed first, particularly those related to the relationships between situated, real-world studies and the adoption of experimentally controlled studies. Criteria for assessment are then presented and discussed.

### 10.4.1 Methodology

The research cuts across many disciplines, namely, industrial design, interaction design, ergonomics, HCI, psychology, social sciences and computing. All provide different focal points and philosophical arguments. For this reason, the early research suffered from what Seago and Dunne (1996) termed 'methodological intimidation'. Methodologies used in these studies were derived from known established academic disciplines with little consideration given to alternative approaches. The early studies were designed to gain design data using *scientifically* formalised and systematic means and were, therefore, positivist in approach. Studies were designed to satisfy predetermined objectives and intentions rather than using the studies as a process of inquiry. Jones (1998) describes this approach as 'utilitarianism'. This is in contrast with the latter studies where methods of investigation used a social constructionism paradigm.

The shift in research inquiry from positivist to social constructionism was somewhat laboured. This was due to an ingrained positivist assumption by the researcher that interaction design data could



successfully be gained as an external corpus of knowledge rather than being embedded within the design process itself. Linked to this, it was assumed that specificity criteria could be satisfied before applicability criteria. The only consolation is that HCI researchers are still making this fundamental mistake by endeavouring to bridge the gap between research and practice by gaining knowledge which can then be suitably packaged and targeted. As Wood (2000 p 45) points out, 'design knowledge is oriented to making practical, appropriate and elegant interventions within actual situations' and that 'factual knowledge is of very limited use without "tacit knowledge"' (p 53). Studies 3 and 4 proved to be important milestones in shifting towards this viewpoint. In these two studies the experiential knowledge gained from carrying out interface design activity was more revealing than their outcomes. On reflection, some form of ethnographically-based study prior to the development of the design tools may have revealed insights that would have accelerated the development process.

Despite this initial methodological intimidation, a variety of research methodologies was used. Design rationale analysis was used in Study 2 (QOC). Formal, experimentally controlled usability trials were carried out in Study 3 using parametric statistical tests. Design data were gathered using less rigorous, 'quick and dirty', methods in Study 4, forming a pre-cursor to the development of the design tools. Different forms of questionnaire surveys were used in Studies 5, 6 and 8 and semi-structured interviews were used in Study 8. Action research studies were carried out within the remaining studies (9 - 12) with content analysis conducted on the raw data produced in Study 11. As far as possible studies were designed to occur in a real-world context.

As the studies in this research attempted to understand or intervene in real-world design activity, they had to be designed around commercial demands and pressures. As McDonnell (1997 p 473) stated, 'the form



and nature of design decision making cannot be gained otherwise than by studying the designing in the context of professional practice'. This requirement inevitably had an effect on the methodological design of the research. Intervening in 'professional practice' required trust on behalf of the participating design organisations. Introducing and evaluating incomplete and unproven design methods further challenged this trust. Such methodological problems were overcome by introducing small, phased, opportunistic studies in which design contingencies could be built into the design process to safeguard real design projects. To minimise the risk to commercial projects, many studies simply had to be conducted 'off-line'. Even with these contingencies, two studies were cancelled at very short notice due to the arrival of a higher priority design project.

Changes in management and organisational restructuring in collaborating organisations meant that the research often had to be 're-sold' to senior management. This was particularly true with Electrolux, who were involved in the research for three years. To gain approval, the commercial benefits of the design tools had to be continually emphasised, such as the marketability of any proposed new interface designs.

Based on this experience, it is perhaps unsurprising that so few academic studies are carried out in unpredictable, real-life commercial contexts.

#### **10.4.2 Assessment criteria**

##### *Integration and coherence*

The thirteen design studies have been presented as closely to their chronological order as possible as the results of one study tended to impact on subsequent studies. The 'thread' that links the studies



together was the effectiveness criteria. Each study examined and tested various elements of the effectiveness criteria, culminating in the Study 12, which examined all the criteria concurrently.

### *Originality and creativity*

This thesis lays claim to originality by considering an area of research that has been neglected. Previous research has not attempted to understand the factors that influence the effectiveness of guidelines using user-centred methods. As the inquiry progressed towards the development of design tools, methods and techniques from the field of HCI, were applied to the field of interaction design. Only one other study was identified throughout the inquiry that had similarly undertaken this approach (Säde 2001). This thesis remains distinct from the Säde work by evaluating this transitional process through the effectiveness criteria. The Säde study was limited to attempting to successfully implement adapted design tools without using objective measures to assess their perceived or actual impact. This research on secondary design methods must not be confused with the abundance of studies that have addressed the effectiveness of *primary* design tools, methods and techniques.

### *Rigour*

From the outset, this research has been pragmatically aimed at developing bridging mechanisms that work effectively in professional interaction design practice. It was important that the outcomes of this research could be put into practice. Grounding practical applications in a real-world context meant the investigative studies were contingent on a wide variety of situated conditions. Rigour was sought through the examination and re-examination of the effectiveness criteria and the use of appropriate methodologies to gain data



## 10.5 Contribution to knowledge

This sections highlights the key contributions to knowledge gained from this inquiry.

*Passive, non-bespoke design guidelines have been shown still to have a minimal affect on interaction design behaviour despite improving their specificity and applicability.*

This finding addressed the initial research question posed Section 1.1 - 'Can effective interface design guidelines be produced for use in the design of future consumer product technologies (CPT)?' The literature review in Chapter 2 provided evidence that existing HCI interface design guidelines were limited in their effectiveness. However, attempts had not been made to specifically develop or improve the effectiveness of generic guidelines using user-centred design methods. Studies 2 and 5 demonstrated that designers find it difficult to frame design problems to fit non-bespoke or externally constructed design guidelines.

*The effectiveness of using interface design data can be improved through interaction designers using secondary design tools if these data are produced and situated within the interaction design problem.*

This finding addressed the revised research question posed in Section 2.8 - 'How can effective interface design data be produced for use in the design of future consumer product technologies (CPT)?' Studies 3 and 4 partially answered this question by demonstrating that design data (and the subsequent design guidelines) were difficult to produce proactively due to a lack of contextually sensitive knowledge that besets novel interaction design problems. The second literature review in Chapter 7 and Studies 9 - 12 answered the research questions more completely and concluded that alternative bridging mechanisms could be developed and adopted by interaction designers, as long as they were in the form of design tools that facilitate experiential learning and



have high applicability and specificity. These studies demonstrated that designers and participants using secondary design tools could effectively use adapted HCI design tools. However, considerable development work is still required in order to improve the organisational survival of these design tools.

*Secondary design tools would benefit from being evaluated by specificity and applicability effectiveness criteria concurrently.*

The effectiveness criteria developed in this research have been shown to be important benchmarking tests for the successful adoption of secondary design tools. These criteria remained robust when applied to this research and other related studies. This inquiry has highlighted that applicability criteria should not be separated from specificity criteria. Both need to be addressed concurrently. The organisational context in which design data is being used is as important as the reliability and validity of the data itself.

## 10.6 Future directions

The findings from this research have sought ways to improve the gap between interaction design research and practice. Many suggestions were offered in the literature review in Chapter 2 about how best this gap could be bridged either through shared team working (Fulton 1993; Haubner 1990; McCelland 1990) or through improving design education (Boff 1998; Gillian and Bias 1991; Rouse *et al* 1991). Evidence from this research has revealed a third alternative, which could be described as a hybrid of the two. Through the use of the design tools, designers can produce their own research-based design data. Design researchers in interaction design need to contribute to this 'third way' by continuing to improve the effectiveness of such tools by carrying out in-the-field evaluation studies of different secondary design tools being used on live design projects. In this way knowledge is produced for the



purposes of improving practical application and not isolated as an area of research interest.

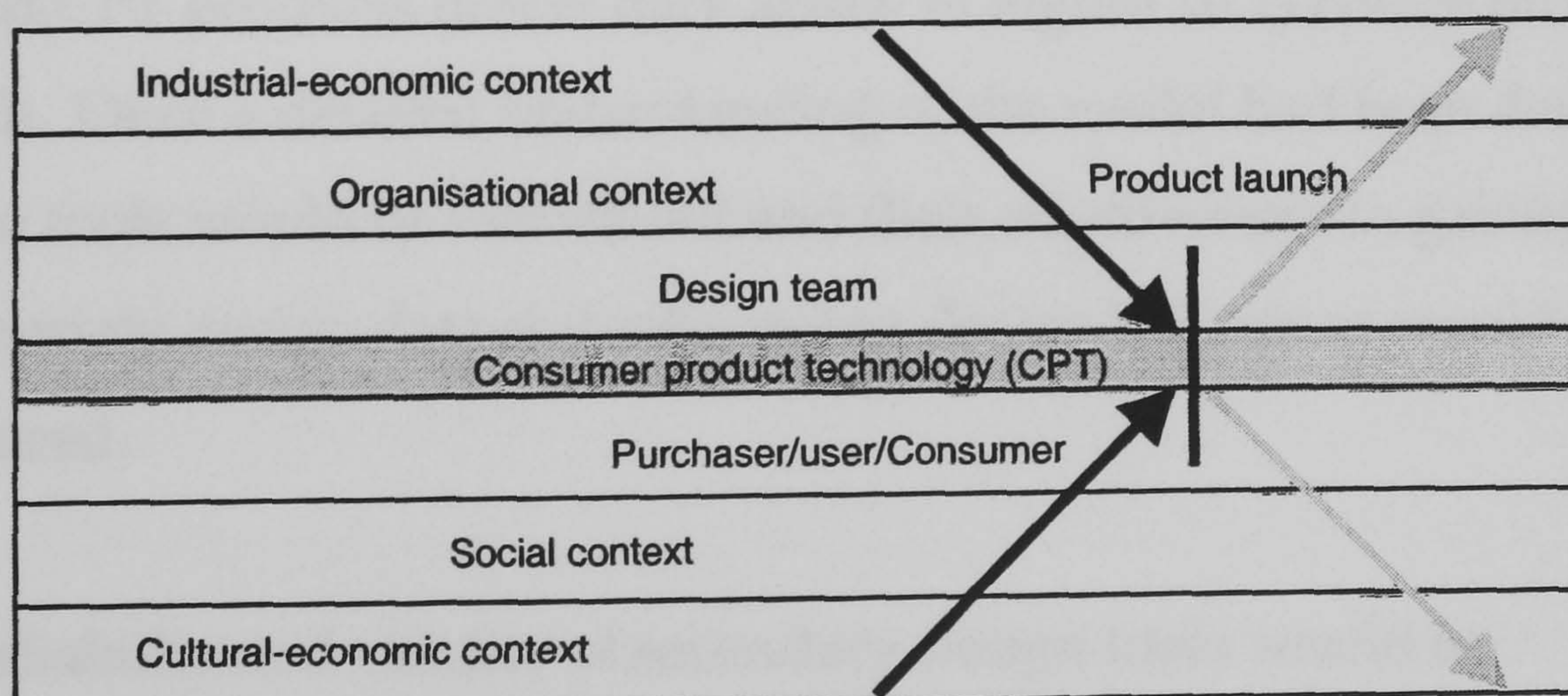
One of the EPSRC reviewers made the following comment about the future direction for this research. 'With this project now completed and the experience gained, the natural progression would be to improve our understanding of the design process itself, of the kind considered in the study'. The assumption embedded in this comment is that interaction design, as a design process, needs to be understood from an external, objective point of view and that once a generalised model has evolved then a better understanding can be gained of how best to apply design tools to that process. This is a positivist viewpoint, whereas findings from this research suggest that the development of design data needs to go hand-in-hand with situated design processes. A social constructionism paradigm would, therefore, appear more productive.

It is true that a deeper understanding of the design process is required but *only* while design tools are being used and evaluated in their situated context. Further research should not be carried out merely to understand an abstracted notion of the interaction 'design process'. Models of the interaction design process should be tested as an integral part of a continuing and iterative development of different forms of prototypical design tools, which are tested against the effectiveness criteria. An example of this type of research is Perry and Sanderson (1998) who examined the pattern of communication and the range and role of artefacts used during the design activity of an engineering and building design group. They found a wide range of informal communication in projects, such as group discussions around drawings. They argued that design tools must reveal their 'location' within the process - when and who has acted upon them. They suggest that current design and communication technologies do not necessarily support existing communication methods and that understanding the



rules of use may improve the design of communication technologies. Building on studies like these could form an important springboard to assessing alternative design tools or variants for gaining effective design data.

Furthermore, the recording, documenting and disseminating of this form of ethnographic research activity also requires attention (Stiff 1998). Different recording, documenting and disseminating methods need to be evaluated, although some of this work has begun such as comparing a reflective practice description (Valkenburg and Dorst 1999) with data-driven descriptions (Mazijoglou *et al* 2000) or through experiential cognition (Benyon and Imaz 1999). Conventional forms of design representation such as sketches, scenarios and storyboards are beginning to be understood in terms of their effectiveness in design (Saddler 2001), but these need to be placed in the broader context of design team usage and organisation survival. Work is required on evolving and devising new forms of accessible interaction models within this context. The further development of the design tools produced in this inquiry should continue while considering this larger frame of organisational context.



**Figure 10.1 Layered communication model**



With these concerns of supporting and recording participative design activity within an organisational context, a study using an investigative communication model is proposed in Figure 10.1. This model presents a layered model of contextual factors that may influence the design of future CPT. Interaction is not defined as being a dialogue between users and a product but is regarded as a series of fragmented dialogues between users, consumer product technologies and designers over a period of time, involving product development, launch and use. Designers and users (the layers closest to the CPT) form and develop design and usage behaviours using knowledge from other parts of the layered model. It is within this context that these and future design tools would be assessed.

An interesting research study would be to track the development of a target CPT product (top dark arrow in Figure 10.1) and to explore the communication points between all layers over the product development period. For example, observations would be made of product development meetings to investigate how product development projects are initiated and factors influencing their direction. Similarly, potential target users would be tracked to observe how social and cultural factors affect their purchase habits and perceptions of similar or related CPT products (lower dark arrow in Figure 10.1) before product launch. Once a detailed understanding of the model had been devised, design tools would be introduced and their effectiveness in gaining appropriate design data and influencing design behaviour would be measured.

The reliability and validity of secondary design tools would be measured by comparing the use of design data derived from such tools against the organisation's natural sources of design data. After product launch, predictions gained from the design tools would be compared against real-world perceptions and usage of the target product (lower



grey arrow) and the perceived effectiveness of the design tools within the design organisation (upper grey arrow). In this way, further improvements to the design tools could be made. These investigations would be achieved through ethnographically-based studies of designers and users in their natural environments, over a period of time that includes product development, launch and use.

The original experimental paradigm used in this research was based on objectivity and controlled variables and proved too restrictive to be useful. Through the findings in this research, a shift in methodological approach involving action research methods was more successful and therefore proposed for future work.

This in-the-field study should reflect many of the methodological characteristics that evolved in this research and used most effectively in Study 12. For future studies to gain effective results, they must be opportunistic and be interventionist. They need to go beyond conventional ethnographic-based studies, which are based on learning and understanding a particular cultural context, and through collaborative means intentionally inject new design methods while monitoring the broader organisational impact that such design tools create.

## 10.7 Final thoughts

Cross (2001 p 55) commented, 'we must avoid swamping our design research with different cultures imported either from the sciences or the arts. This does not mean that we should completely ignore these other cultures. On the contrary, they have much stronger histories of inquiry, scholarship, and research than we have in design. We need to draw upon those histories and traditions where appropriate, while building our own intellectual culture, acceptable and defensible in the world on its own terms.' This 'intellectual culture' has to be rooted in experiential



learning. Design researchers need to ensure that new design data emerges *through* experiential learning. This would help move beyond the still typical approaches to interaction design research and practice. The way design data are gained, implemented and disseminated requires a quite different design research landscape. The effectiveness criteria, used in this thesis, set out methodological benchmarks by which future design tools should be measured. In this way interaction design will begin to form its own research culture - out in the vegetable patch.



# 11 References

---

- Adler, P. S. and Winograd, T. A., 1992, The usability challenge. In *Usability: Turning Technologies into Tools* (eds) P S Adler and T A Winograd Oxford University Press, New York.
- Alben, L., Faris J. and Saddler H., 1994 Making it Macintosh, *Interactions*, 1, (1) pp10-20
- Alben, L., 1996 Defining the criteria for effective interaction design, *Interactions*. 3 (3) pp11 - 16.
- Andre, T., Harston, H. R., Belz, S. M., and McCreary, F. A., 2001, The user-action framework: a reliable foundation for usability engineering support tools, *International Journal of Human-Computer Studies*, 54 (1) pp107 - 136.
- Bailey, R. W., 1982, *Human Performance Engineering: A guide for system designers*. Prentice Hall, Inc. New Jersey.
- Ball, L.J. and Ormerod, T. C. 2000, Putting ethnography to work: the case for a cognitive ethnography of design, *International Journal of Human-Computer Studies*. 53 (1) pp147 - 168
- Barnard, P., 1991 Bridging between basic theory and the artifacts of Human -Computer Interaction. In Carroll, J. M., (ed). *Designing Interaction: Psychology at the Human-Computer Interface*, Cambridge, Cambridge University Press.
- Barnard, P. May J., Duke, D. and Duce, D., 2001 Systems, Interactions, and Macrotheory, *ACM Transactions on Computer-Human Interaction*, 7, (2) pp 222 - 262.



- Bastien, J. M. and Scapin D. L., 1995a, Evaluating a user interface with ergonomics criteria, *International Journal of Human Computer Interaction* . 7 (2) pp 105-121.
- Bastien, J. M. and Scapin, D. L., 1995b, How usable are usability principles, criteria and standards? *Symbiosis of Human and Artefact: Future Computing and Design for HCI*, Anzai, F., Ogawa, H. and Mori, H. (eds), Elsevier, Amsterdam, pp 343-348
- Benyon, D. and Imaz, M., 1999, Metaphors and models: conceptual foundations of representation in interactive systems development, *Human Computer Interaction*, 14 (1,2) pp159 - 189.
- Bergman, E., 2000, *Information appliances and beyond*, Academic Press, San Diego.
- Bevan, N., 2001 International standards for HCI and usability, *International Journal of Human-Computer Studies*, 55 (4) pp533 - 552
- Beyer, H. and Holtzblatt, K., 1998 *Contextual Design: Defining Customer-Centred Systems*, Morgan Kaufmann, San Francisco, California.
- Bias, R. G., 1994, *Cost-justifying usability*, Academic Press, Boston.
- Bishop, E. W., and Guinness, G. V., 1966, Human factors in interaction with industrial design, *Human Factors*, 8 pp 279 - 289
- Black, A. and Buur J., 1996 'Making solid user interfaces work' *Information Design Journal*, 8 (2) pp 99-108
- Blattner, M. M., Sumikaira, D. A., and Greenberg, R. M., 1989, Earcons and Icons: their structure and common design principles, *Human Computer Interaction*, 4 pp 11-44
- Bødker S., 1998, Understanding representation in design, *Human Computer Interaction*, 13 (2) pp107 - 125.



- Bødker S., 2000, Scenarios in user-centred design - setting the stage for reflection and action, *Interacting with Computers*, 13 (1) pp61 - 75.
- Bødker, S., 1991 *Through the interface: a human approach to user interface design*. Lawrence Erlbaum, New Jersey.
- Bødker, S., 1996, Applying Activity Theory to Video Analysis: How to make Sense of Video Data in Human-Computer Interaction. In Nardi, B A (ed) *Context and Consciousness*, MIT Press, Cambridge, Massachusetts.
- Boff, K. R., (1988) The value of research is in the eye of the beholder, *Human Factors Bulletin*, 31 (6) pp 1-4
- Bolt, R A and Herranz, E. 1992, Two handed gesture in multi-modal natural dialogue, *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, Monterey, California, pp 7-14
- Bonner, J. V. H., 1998, Towards Interface Design Guidelines for Consumer Products in Stanton N (ed) *Human Factors in Consumer Products*, Taylor Francis, London.
- Bonner, J. V. H. and van Schaik P., 1998, The use of high and low level prototyping methods for product user interfaces, In *Proceedings of the Annual Ergonomics Conference, Royal Agricultural College, Cirencester*, 1-3 April 1998
- Bonner, J. V. H., 1997, Developing advanced interface design guidelines from survey based and empirical research, In *Proceedings of the International Conference on Design and Technology Education, Research and Curriculum Development*, Loughborough University.
- Bonner, J. V. H., 1999a, *Implications for using intelligence in consumer products*, In Green W. S. and Jordan P. W., (eds), *Human Factors in Product Design: Current Practice and Future Trends*, Taylor and Francis, London.



- Bonner, J. V.H., 1999b Exploring the 'interface culture' between ergonomics and product interface design, *Proceedings of the European Academy of Design: Design Cultures Conference*, Sheffield Hallam University, 03 March - 1 April 1999.
- Brooks, R., 1991, Comparative task analysis: An alternative direction for human-computer interaction science. In J. Carroll, (ed) *Design Interaction: Psychology at the human Computer Interface*, Cambridge: Cambridge University Press.
- Brown, J. S. and Duguid, P., 1992, *Enacting design for the workplace*. In *Usability: Turning Technologies into Tools* Eds. P S Adler and T A Winograd Oxford University Press, New York.
- Brown, J. S. and Duguid, P., 1994 Borderline Issues: Social and Material Aspects of Design, *Human-Computer Interaction* 9 (1) pp 3-36
- Brown, B. A. and Perry M., 2000, Why don't telephones have off switches? Understanding the use of everyday technologies a research note, *Interacting with Computers*, 12 (6) pp623 - 634.
- Brown, C. M., 1988, *Human-Computer Interface Design Guidelines*, Ablex Publishing
- Burford, B C and Baber C, 1994, A user-centred evaluation of a simulated adaptive autoteller, *Contemporary Ergonomics 1994* Robertson S.A.(ed), Taylor and Francis, London pp. 46-51
- Burns, C.M., Vicente J.M., Christoffersen K. and Pawlak W. S., (1997) Towards viable, useful and usable human factors design guidance, *Applied Ergonomics*, 28 (5/6) pp 311-322
- Button, G., 2000, The ethnographic tradition and design, *Design Studies*, 21 (4) pp319 - 332.
- Buur, J., Bagger, K. and Binder, T., 1997, Turning usability dialogue into user dialogue, *Proceedings of the 13th Triennial Congress of the International Ergonomics Association*, Tampere, Finland.



- Candy, L., and Edmonds E. A., 1997, Supporting the creative user: a criteria-based approach to interaction design, *Design Studies*, 18 pp 185 - 194
- Carroll, J. M., 1995, The Scenario perspective on system development, In Carroll JM (ed) *Scenario-based design: envisioning work and technology in system development*, John Wiley and Son, New York.
- Carroll, J. M., 1991 *Designing Interaction: Psychology at the Human-Computer Interface*, J. M. Carroll (ed) Cambridge University Press.
- Carroll, J. M., 2000, Five reasons for scenario-based design, *Interacting with Computers*, 13 (1) pp 43 - 60.
- Carter, J., 1999, Incorporating standards and guidelines in an approach that balances concerns for developers and end users, *Interacting with Computers*, 12 (2) pp 179 - 206.
- Clark, T.S. and Corlett E. N., 1984, *The ergonomics of workspaces and machines: A design manual*, Taylor and Francis, London.
- Conklin, J. E. and Burgess Y., 1991, A process-oriented approach to design rationale, *Human Computer Interaction*. 6 (3/4) pp 357-391
- Cross, N., 2001, Designerly ways of knowing: Design discipline versus design science, *Design Issues*, 17 (3) pp 49 - 55
- Cuomo P. L. and Bowen, C. D., 1994 Understanding usability issues addressed by three user-system interface evaluation techniques, *Interacting with Computers*, 6 (1) pp 86-108
- Cushman, W. H. & Rosenberg, D. J., 1991, *Human Factors in Product Design*, Elsevier, Amsterdam.
- De Souza, F L, Long, J B and Bevan, N., 1990, Types of error and difficulty in using human-factors guidelines: the case of interface menu design, *Contemporary Ergonomics*, 1990, Lovesey E.J. (ed) Taylor and Francis, London. pp. 340-346



- Denning, P. and Dargan, P. (1996), Action centred design. In Winograd T (ed), *Bringing design to software*, Addison-Wesley, Massachusetts.
- Diffrient, N., Tilley, A. R. and Bardagjy, J. C., 1985 *Humanscale 1/2/3*, The MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142
- Directory of Designers*, 1991 The Design Council, London
- Dix, A., Finlay, J., Abowd, G. and Beale, R., 1998 *Human-Computer Interaction*, Second Edition, Prentice Hall Europe, London.
- Earthy, J., Sherwood Jones, B. and Bevan, N., 2001, The improvement of human-centred processes - facing the challenge and reaping the benefit of ISO 13407, *International Journal of Human-Computer Studies*, 55 (4) pp553 - 585
- Ehn, P., 1992, Scandinavian Design: On Participation and Skill. In Adler P S and Winograd T A (eds) *Usability: Turning Technologies into Tools*. Oxford University Press, New York.
- Fields, R., and Wright P., 2000, Editorial: Understanding work and designing artefacts, *International Journal of Human Computer Studies*, 53 (1) pp1 - 4.
- Fitzmaurice, G. W., Ishii, H. and Buxton, W., 1995, Bricks: laying the foundations for graspable user interfaces, In *CHI '95 - Mosaic of Creativity*, Proceedings of Conference on Human Factors in Computer Systems, Denver, Colorado, pp 442-449
- Fleck, J., Molina, A. and Nicoll, D., 1997 *Designing smart products*. Paper presented to Design council Research Workshop, London, June/July 1997.
- Fowler, S. and Stanwick, V., 1994, *The G U I style guide*, Academic Press Professional



- Frolich, D. M., and Dray S., 2001, Breaking up is hard to do: Family perspectives on the future of the home PC, *International Journal of Human-Computer Studies*, 54 (5) pp701 - 724.
- Fulton, J., 1993, Users trials for information graphics: replacing designers' assumptions with feedback from users. *Information Design Journal*, 7 (2) pp 99 - 180
- Galer, I., (ed) 1987, *Applied Ergonomics Handbook*, Second Edition. Butterworth and Co.
- Galitz, W. O., 1993, *The essential guide to user interface design*, QED Information Sciences, Boston
- Gardiner, M. M. and Christie, B., 1987, *Applying cognitive psychology to user-interface design*, John Wiley and Sons, Chichester.
- Gaver, W. W., 1989 The Sonic Finder: An interface that uses auditory icons. *Human Computer Interaction*, 4 pp 67-94
- Gillan, D. J. and Bias, R. G., 1992, The interface between human factors and design, In *Proceedings of the Human Factors Society 36th Annual Meeting*, 1 pp 443 - 447
- Gittens, D., 1986, Icon based human computer interaction, *International Journal of Man Machine Studies*, 24 pp 519-543
- Gould, J. D., Boies S. J. and Lewis C., 1991, Making usable, useful productivity, *Communications of the ACM*, 34 (1) pp 75-85
- Gould, J. D., Lewis, C. H., 1985, Designing for Usability - Key Principles and What Designers Think, *Communications of the ACM*, 28 pp 300 - 311
- Grandjean, E., 1988, *Fitting the task to the man: A textbook of occupational ergonomics*, Fourth Edition. Taylor and Francis.
- Grant, J. and Fox, F., 1992, Understanding the role of the designer in society. *Journal of Art and Design Education*, 11 1.



- Green, T. R. G. and Petre, M., 1996, Usability analysis of visual programming environments: a 'cognitive dimensions' framework. *Journal of Visual Languages and Computing*, 7, pp131-174.
- Greenbaum, J and Kyng M 1991 *Design at Work*. Hillsdale, NJ: Lawrence Erlbaum Associates
- Grudin, J., 1991, Systematic sources of suboptimal interface design in large product development organisations, *Human-Computer Interaction*, 6 pp147 - 196
- Grudin, J., and Poltrock, S. E., 1989, User interface design in large corporations: co-ordination and communication across disciplines, *Proceedings of ACM CHI'89, Conference on Human Factors in Computing Systems*, pp 197-203
- Halstead-Nussloch, R., 1989, The Design of Phone-Based Interfaces for Consumers, *Proceedings of ACM CHI'89 Conference on Human Factors in Computing Systems*, pp 347-352
- Hammond. N., Jorgensen, A., MacLean, A., Barnard, P., & Long, J., 1983, Design Practice and Interface Usability: Evidence from Interviews with Designers, *Proceedings of CHI' 83 Conference, Boston, Mass, 12 -15 December 1983*, North Holland, Amsterdam, pp 40 - 44.
- Haubner, P. J., 1990, Ergonomics in industrial product design, *Ergonomics*, 33 (4) pp 477-485
- Helander, M., 1995, *A guide to the ergonomics of manufacturing* Taylor & Francis, London.
- Helander, M., (ed) 1988, *Handbook of Human-Computer Interaction*, North-Holland, Amsterdam.
- Henninger S, Haynes, K. and Reith M.W., 1995, A framework for developing experienced-based usability guidelines, *Proceedings of*



- the Symposium on Designing Interactive Systems, DIS95 Ann Arbor*  
pp 45-53
- Henninger, J., 2000, A methodology and tools for applying context specific usability guidance to interface design, *Interacting with Computers*, 12 (3) pp 225 - 243
- Hill, T., Smith, N.D. and Mann, M.F., 1987, *Journal of Applied Psychology*, 72 (2), 307-313.
- Hix, D and Hartson H R 1993, *Developing user interfaces: ensuring usability through product and process*, J Wiley and Sons, New York
- Hollnagel, E. and Woods, D. D., 1983, Cognitive systems engineering: new wine in new bottles, *International Journal of Man-Machine Systems*, 18 pp583 - 600.
- Hubona, G.S. and Blanton, J.E., 1996, Evaluating system design features, *International Journal of Human-Computer Studies*, 44 pp 93-118.
- Hughes, J A Randell D. and Shapiro D., 1996, *Faltering from Ethnography to Design In Wise J A., Hopkin D V and Stager P., (eds) Verification and Validation of Complex Systems: Additional Human Factors Issues*, Embry Riddle Aeronautical University Press, Florida.
- Hughes, J. A., King, V. Rodden, T. and Anderson H., 1995, The role of Ethnography in Interactive Systems Design, *Interactions* 11 (2) pp 56-65
- Hutchins, E., 1994, *Cognition in the Wild*. Cambridge, MA: MIT Press
- Jacob, J. K., Leggett J. J., Myers, B. A., and Pausch. R. 1993, Interaction styles and input/output devices, *Behaviour & Information Technology*, 12 (2) pp 69-79
- Jacobson, J., 2001, The desktop Fab, *Communications of the ACM*, 44 (3) pp41 - 42



- Jeffries, R, Miller, J R, Wharton, C and Uyeda, K M., 1991 User interface evaluation in the real world: a comparison of four techniques, Reaching through technology. In Robertson, S. P., Olson, G. M. & Olson, J. S. (eds) *Proceedings of CHI'91*, New Orleans, USA, ACM Press, pp 119-124
- Johnson, C., 1996, What is a PhD in HCI? In Blandford A and Thimbleby H., (eds) *HCI '96 Industry Day and Adjunct Proceedings of HCI Conference*, London, 20 - 23 August, 1996.
- Jones, J. C. and Jacobs, D. 1998, PhD research in design, *Design Studies*, 19 (1) pp5-7
- Jordan, B and Henderson A., 1995, Interaction analysis: Foundations and Practice, *The Journal of the Learning Sciences* 4 (1) pp 39-103
- Junestrand, S., Keijer, U. and Tollmark, K., 2001, Private and public digital domestic spaces, *International Journal of Human-Computer Studies*, 54 (5) pp 753 - 778.
- Kaptelinin, V., 1996 Computer mediated activity, In Nardi, B. A., (ed) *Context and Consciousness*, MIT Press, Cambridge, Massachusetts.
- Karat, J., 1995 Scenario use in the design of speech recognition system, In Carroll, J. M., (ed) *Scenario-based design: envisioning work and technology in system development*, John Wiley and Sons, New York pp109 - 134
- Kelly, T., 2001 Prototyping is the shorthand of innovation, *Design Management Journal*, 12 (3) pp35 - 42.
- Kemp, J.A.M., and van Gelderen T., 1993, The co-discovery method: an informal method for iteratively designing consumer products, *IPO Annual Progress Report 28*, Institute for Perception and Research, Eindhoven, The Netherlands, pp 143-150
- Klein, G. A. and Brezovic, C. P., 1986 Design engineers and the design process: Decision strategies and human factors literature,



- Proceedings of the Human Factors Society 30th Annual Meeting, 29th Sept - 3rd Oct Dayton Ohio, The Human Factors Society, Santa Monica, California, pp 771-775*
- Krippendorff, K., 1980, *Content Analysis: An Introduction to Its Methodology*, Sage Publications, Beverly Hills.
- Kuhn, S., 1996, Design for people at work. In Winograd T (ed), *Bringing design to software*, Addison-Wesley, Massachusetts.
- Kukla, C. D., Clemens, E. A., Morse, R. S., and Cash D., 1992, Designing effective systems: A tool approach. In Adler P. S. and Winograd T. A., (eds) *Usability: turning technologies into tools*, Oxford University Press, New York.
- Kuutti, K 1996, A framework for HCI research, In Nardi, B A (ed) *Context and Consciousness*, MIT Press, Cambridge, Massachusetts.
- Kwahk, J., Han, S. H., Yun, M. H., Hong, S. W., Chung, M. K., and Lee, K. S., 1997, Selection and classification of usability attributes for evaluating consumer electronic products, *Proceedings of the Human Factors and Ergonomics Society 41<sup>st</sup> Annual Meeting*, Albuquerque.
- Kyng, M., 1995, Creating Contexts for Design, In Carroll, J. M., (ed) *Scenario-based design: envisioning work and technology in system development*, John Wiley and Son, New York.
- Lafrenière, D., 1996, CUTA: A simple, Practical and Low-cost Approach to Task Analysis, *Interactions*, 3 (5) pp 35-39
- Laurel, B 1993, *Computer as Theatre*, Addison-Wesley, Reading Massachusetts
- Lavery, D., Cockton G. and Atkinson M. P., 1997, Comparison of evaluation methods using structured usability problems, *Behaviour and Information Technology* 16 (4/5) pp 246-266
- Lawson, B., 1980, *How designers think*, Architectural Press, London, UK.



- Leontjev, A. N., 1981, *Problems of the Development of Mind*, Moscow Progress.
- Lewis, S., Mateas M., Palmiter S. and Lynch, G., 1996 Ethnographic data for product development: A collaborative process. *Interactions* 3 (6) November/December pp52 - 69
- Lu, S., Yoshizaka, S., Kamiya, T. and Miyai, H., 1995, A human-computer dialogue agent with body gestures, hand motion and speech, In *Symbiosis of Human and Artifact: Future Computing and Design for HCI*, Anzai, F., Ogawa, H. and Mori, H. (eds) Elsevier, Amsterdam, 1 pp 17-22
- Macaulay, C., Benyon, D. and Alison C, 2000, Ethnography, theory and systems design: from intuition to insight, *International Journal of Human-Computer Studies*, 53 (1) pp35 - 60.
- Mack, R. L. 1995, Scenarios as Engines of Design, In Carroll, J.M. (ed) *Scenario-based design: envisioning work and technology in system development*, John Wiley and Son, New York.
- MacLean A and McKerlie D., 1995, Design space analysis and use representation, In Carroll, J.M. (ed) *Scenario-based design: envisioning work and technology in system development*, John Wiley and Son, New York.
- Maes, P., 1995, Artificial life meets entertainment: Lifelike autonomous agents, Special Issue on New Horizons of Commercial and Industrial AI, *Communications of the ACM*, 38 (11)
- Maguire, M. C., 1999, A review of user-interface design guidelines for public information kiosk systems, *International Journal of Human-Computer Studies*, 50 (3) pp263 - 286.
- Martin, A. and Eastman, D., 1996, *The user interface design book*, John Wiley and Sons, Chichester.



- Mazijoglou, M., S.A.R. Scrivener and Valkenburg, R. C., 2000, Matching descriptions of team design, In Scrivener, S. A. R., Ball, L. J. and Woodcock, A., (eds) *Collaborative Design, Proceedings of CoDesigning 2000*, Springer, London.
- McClelland, I., 1990, Marketing ergonomics to industrial designers, *Ergonomics*, 33 (4) pp 391-398.
- McDonnell, J., 1997 Descriptive models for interpreting design, *Design Studies*, 18 (4) pp 457 - 473.
- Monk, A., Nardi B., Gilbert N., Mantei, and McCarthy J., 1993, Mixing Oil and water? Ethnography versus experimental psychology in the study of computer-mediated communication. *INTERCHI '93 Proceedings of Conference on Human Factors in Computing Systems*, Amsterdam, The Netherlands 24 - 29 April 1993.
- Monk, A., Wright P., Haber, J. and Davenport L., 1993, *Improving your Human-Computer Interface: A Practical Technique*. New York, Prentice Hall.
- Mosier, J. N., and Smith, S. L., 1986, Application of guidelines for designing user interface software, *Behaviour and Information Technology*, 5 pp 39-46
- Mountford, S., J. and Johnson, B., 1997, Educational challenges for the new interactivity: Interval research's university workshop, *Design Management Journal*, 8 (3) pp51 - 57.
- Muller, M. J., Tudor, L. G., Wildman, D. M., White, E. A., Root, R. W., Dayton, T., Carr, R., Diekmann, B., and Dykstra-Erickson, E. A., (1995). Bifocal tools for Scenarios and Representations in participatory activities with users. In Carroll JM (ed) *Scenario-based design: envisioning work and technology in system development*, John Wiley and Son, New York.



- Nardi, B. A., 1996 Some reflections on the Application of Activity Theory. In Nardi, B A (ed) *Context and Consciousness*, MIT Press, Cambridge, Massachusetts.
- Nardi, B. A., 1996 Studying context. In Nardi, B A (ed) *Context and Consciousness*, MIT Press, Cambridge, Massachusetts.
- Newman, W. M. and Lamming, M. G., 1995, *Interactive system design*, Addison Wesley Reading UK
- Nielsen, J., 1993 *Usability Engineering*, AP Professional, Boston.
- Nielsen, J., 1994, Heuristic Evaluation, In Nielsen, J. and Mack, R. L., (eds) *Usability Inspection Methods*, John Wiley, New York.
- Nomura, T. and Brownlow, M, 1996 Thinking microwave oven - Logicook *Proceedings of the IEE Artificial Intelligence in Consumer and Domestic Products Colloquium*, 22 October, 1996 Savoy Place, London.
- Norman, D. A., 1991, Cognitive artifacts. In J. Carroll, (ed) *Designing Interaction: Psychology at the human Computer Interface*, Cambridge: Cambridge University Press.
- Norman, D. A., 1986, Cognitive Engineering. In Norman D. A. and Draper S. W, (eds) *User Cenered System Design: New Perspectives on Human-Computer Interaction*, pp31-61. Hillsdale, N J: Lawrence Erlbaum Associates.
- Norman, D. A., 1999, *The invisible computer*, MIT Press, Cambridge, Massachusetts.
- Owen, C. L., 1998, Design Research: building the knowledge base, *Design Studies* 19 pp9-20
- Pare, G., Elam, J.J. 1996, Discretionary use of personal computers by knowledge workers: testing a social psychology theoretical model. *Behaviour and Information Technology*, 14 (4), pp 215-228.



- Pedersen, J. and Buur, J., 2000, Games and movies: towards innovative co-design with users. In Scrivener, S. A. R., Ball, L. J. and Woodcock, A. (eds) *Collaborative Design, Proceedings of CoDesigning 2000*, Springer, London.
- Perry, M., and Sanderson, D., 1998, Coordinating joint design work: the role of communication and artefacts, *Design Studies*, 19 (3) pp 273-288.
- Pheasant, S., 1987 *Ergonomics standards and guidelines for designers*. British Standards Institute, Milton Keynes.
- Piernot, P.P., Felciano R.M., Stancel R., Marsh J., and Yvon M., 1995, Designing the PenPal: Blending hardware and software in a user-interface for children, *CHI '95 - Mosaic of Creativity, Procedures of Conference on Human Factors in Computer Systems*, Denver, Colorado, pp 511-518
- Polson, P., Lewis, C., Rieman, J. and Wharton, C. 1994, Cognitive Walkthroughs: A method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36 pp741-773.
- Poole. S. and Simon M., 1997, Technology trends, product design and the environment, *Design Studies*, 18 (3) pp237 - 248
- Potter, R. Berman, M., Shneiderman B., 1989 An experimental evaluation of three touch screen strategies within a hypertext database, *International Journal of Human Computer Interaction*, 1 pp 41-52
- Preece, J., Rogers Y. and Sharp, H., 2002, *Interaction Design: Beyond Human-Computer Interaction*, John Wiley, Chichester.
- Ravden, S, J., and Johnson, G, I., 1989, *Evaluating usability of human-computer interfaces: a practical method*, Ellis Horwood Ltd, Chichester.



- Rivlen, C., Lewis, R. and Davies-Cooper, R., 1990, *Guidelines for screen design*, Blackwell Scientific Publications, Oxford
- Rogers, Y. and Bellotti V., 1997, Grounding blue sky research: How can ethnography help? *Interactions*, 4 (3) May and June 1997 pp58 - 63
- Rouse, W. B. and Boff, K. R., 1998, Packaging Human Factors for designers, *Ergonomics in Design*, 6 (1)
- Rouse, W. B., 1991, *Design for success - A human centred approach to designing successful products and systems*. John Wiley & Sons, Chichester.
- Rouse, W. B., Cody W. R., and Boff, K. R., 1991, The human factors of system design: understanding and enhancing the role of human factors engineering, *International Journal of Human Factors in Engineering*, 7, pp 87-104
- Rudd, J., Stern, K., Scott, I., 1996, Low High Fidelity Prototyping Debate, *Interactions*, 3 (1) January pp76 - 85
- Saddler, H. J., 2001, Understanding design representation, *Interactions*, 8 (4) pp17 - 23
- Säde, S., 2001, User-centred design: a method in product design consultancy, *The Design Journal*, 4 (3) pp20 - 31
- Sakamura, K., 1993, *TRON Human-Machine Interface Specification*, TRON Association, Japan
- Sanders, M. S. and McCormick, E. J., 1992, *Human factors in Engineering and design* 7th edition Maidenhead Berkshire, McGraw-Hill
- Sanders, E B -N, 2000, Generative tools for co-designing. In Scrivener S A R, Ball L J and Woodcock A (eds), *Collaborative Design, Proceedings of CoDesigning 2000*, Springer, London.



- Scapin, D. L. and Bastien, J. M., 1997, Ergonomic criteria for evaluating the ergonomic quality of interactive systems, *Behaviour and Information Technology*, **16**, (4/5) pp220 - 231
- Schön, D. A., 1983, *The Reflective Practitioner*, Basic Books
- Schrage M., 2000 Serious play: the future of prototyping and prototyping the future, *Design Management Journal* **11** (3) pp50 - 57.
- Schuler, D and Namioka A, (eds) 1993, *Participatory design: principles and practices*, Hillsdale, NJ: Lawrence Erlbaum Associates
- Schumacher, R. M, Hardzinski, M. L. and Schwartz, A. L., 1995, Increasing the usability of interactive voice response systems: research and guidelines for phone-based interfaces, *Human Factors*, **37** (2) pp 251-264
- Seago, A. and Dunne A., 1996/7 New Methodologies in art and design research: The object is discourse, Royal College of Art research papers, **2** (1) pp 1 - 5
- Sears A and Shneiderman, B., 1991, High precision touchscreens: design strategies and comparisons with a mouse, *International Journal of Man Machine Studies* **34** pp 593-613
- Sears, A Revis, D., Swatski, J., Crittenden, R. and Shniederman, B., 1993, Investigating touchscreen typing: the effect of keyboard size on typing speed, *Behaviour and Information Technology*, **12** (1) pp17-22
- Shedroff, N, 1999, Information Interaction Design: A unified field theory of design, In Jacobson R, (ed) *Information design*, MIT Press Cambridge Massachusetts.
- Shimada, N, Shirai, Y and Kuno, Y., 1995, Hand gesture recognition using computer based vision based on model-matching method, In *Symbiosis of Human and Artifact: Future Computing and Design*



- for HCI*, by Anzai F, Ogawa, H. & Mori, H. (eds), Elsevier, Amsterdam, pp. 11-16
- Shneiderman, B., 1992, *Designing the user interface: strategies for effective human computer interaction*, 2nd edition, Reading MA, Addison Wesley
- Sidhu, C. K., Coyle, G., 1995, Delivering the promise of speech interfaces, In *Symbiosis of human and artifact: Future computing and design for HCI*, Anzai, F., Ogawa, H. and Mori, H, (eds) Elsevier, Amsterdam pp529-533
- Smith, S. L., and Mosier, J.N. 1984, Design guidelines for user-system interface software, *Technical report ESD-TR-84-190*, USAF Electronic Systems Division, Hanscom Air Force Base, Massachusetts. NTIS No AD A154 907
- Stiff, P., 1998, Publishing Design Research: Miscounting Quality, *The Design Journal*, 1 (3) pp51 - 60.
- Stumpf, W., 1984 A serious chair: Observations and Intentions, *Design Quarterly*, 126, p35.
- Suchman, L. 1987, *Plans and situated Actions*, Cambridge: Cambridge University Press
- Suri, J., F. and Marsh, M., 1997, Human Factors in design - Suggestions for effective communication in multi-disciplinary teams, *Proceedings of the Human Factors and Ergonomics Society 41<sup>st</sup> Annual Meeting*, Albuquerque.
- Sutcliffe, A. G., and Carroll, J. M., 1999, Designing claims for re-use in interactive systems design, *International Journal of Human-Computer Studies*, 50 (3) pp213 - 241.
- Teasley, B., Scholtz J., 1997 User-Centered GUI Standards Design: A Case Study, In Howard, S., Hammond J., Lindgaard G., (eds) *Proceedings of Human-Computer Interaction Interact '97 IFIP TC13*



- International Conference on Human-Computer Interaction 14th - 18th July 1997*, Sydney, Australia, Chapman and Hall, pp 23-3
- Tetzlaff, L and Schwartz, D R., 1991, The use of guidelines in interface design, Reaching through technology. In Robertson, S. P., Olson, G. M. & Olson, J.S. (eds) *Proceedings of CHI'91*, New Orleans, USA, ACM Press, pp 329-333
- Thomas, B., (1996) 'Quick and Dirty' Usability Tests, In Jordan, P. W., Thomas, B., Weerdmeester, B. A., and McClelland, I. L. (eds), *Usability evaluation in Industry*, Taylor and Francis Ltd, London.
- Trenner, L and Bawa, J., 1998, *The politics of usability*, Springer, London.
- Turner, P. and Turner, S., 1997, *Using Activity Theory to Structure Scenarios for Systems*, Unpublished paper
- Valkenburg, R., K Dorst., 1998, The reflective practice of design teams, *Design Studies* 18 (3) pp 249 - 271.
- Vanderdonckt, J., 1999, Development milestones towards a tools for working with guidelines, *Interacting with Computers*, 12 (2) pp81 - 118.
- Veer, G. C., van der 1990, *Human-computer interaction: learning, individual differences, and design recommendations*. Alblasterdam: Haveka
- Viller, S. and Sommerville, I., 1999, Coherence: An approach to representing ethnographic analysis in systems design, *Human Computer Interaction*, 14 (1/2) pp 9 - 41.
- Vygotsky, L. S., 1978, *Mind and Society*, Cambridge MA: Harvard University Press
- Webb, B. R., 1996, The role of users in interactive systems design: when computers are theatre do we want the audience to write the script? *Behaviour and Information Technology* 15 (2) pp 76-83



- Weed, B., 1996, Visual Interaction Design: The industrial design of the software industry, *SIGCHI Bulletin*, 28 (3).
- Weilenmann, A., 2001, Negotiating use: making sense of mobile technology, *Personal and Ubiquitous Computing*, 2 pp137 - 145.
- Wharton, C., Rieman, J., Lewis, C., and Polson, P., 1994, The cognitive walkthrough: A practitioner's guide, In *Usability Inspection Methods*. John Wiley, New York.
- Wickens, C. D. and Andre, A. D., 1988, Proximity compatibility and the object display, *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA. pp 1335-1339
- Wiklund, M., 1994, *Usability in Practice: How companies develop user-friendly products*, Academic Press, London.
- Wilson, J. and Corlett, E. N., 1990, *Evaluation of Human Work: A Practical Ergonomics Methodology*, Taylor and Francis
- Winograd, T (ed) 1996, *Bringing design to software*, Addison-Wesley, Massachusetts.
- Winograd, T., 1994 Designing the designer. Special Issue on Context in Design, *Human-Computer Interaction*, 9 (1) pp 128 - 132
- Winograd, T. and Flores, S., 1986, *Understanding Computers and Cognition*, Norwood, NJ: Ablex.
- Wixon, D., 1995, Qualitative research methods in design and development, *Interactions* Oct pp 19-26
- Wood, J., 2000, The culture of academic rigour: does design research really need it? *The Design Journal* 3 (1) pp 44 - 57
- Wood, W. T., and Wood, S. K. 1987, Icons in everyday life. In G. Salvendy, S. L. Sauter, and J. J. Sauter, and J. J. Hurrell, Jr. (eds), *Social ergonomic and stress aspects of work with computers*, Amsterdam: Elsevier, pp 97 - 104



- Woods, D. D., Patterson, E. S., Corban, J. M., Watts J. C., 1996, Bridging the Gap Between User-Centered Intentions and Actual Design Practice, *Proceedings of the Human Factors and Ergonomics Society 40<sup>th</sup> Annual Meeting Philadelphia, Pennsylvania, September 2-6, 1996*. The Human Factors and Ergonomics Society, Santa Monica, California, 2 pp 967-971
- Woodson, W. E., Tillman, B. and Tillman, P., 1992, *Human factors in design handbook: information and guidelines for the design of systems, facilities, equipment and products for human use*, 2nd edition New York: McGraw-Hill
- Zimmerman, T. G., Smith, J.R., Paradiso, J.A., Allport, D. and Gershenfeld, N., 1995, Applying electric field sensing to human computer interaction, In: *CHI '95 - Mosaic of Creativity*, Procedures of Conference on Human Factors in Computer Systems, Denver, Colorado, 7-11 May 1995, pp. 280-287
- Zuber-Skerrit, O., 1992, *Action Research in Higher Education*, Kogan Page, London.



# 12 Acknowledgements

---

Much of this research was conducted in close collaboration with industrial design groups. In particular, the design group at Electrolux provided the most significant contribution. I would like to thank the following for their support: Sean Carney (Industrial Design Manager); Phillip Thompson (Senior Industrial Designer); Jonathan Dalton, Darren West, Pelle Ostlund, John Sproats, Paul Flowers, (Industrial Designers who worked on design tool development studies); Graham Drage, (Product Development Manager), and Les Wicks (Marketing Manager) who provided an important business case for this research.

I would also like to thank the other design organisations who contributed: Ian Ferris the Design Manager at Raychem; Members of the User-Centred Design groups at Danfoss - Jacob Buur (Group Manager:), Thomas Binder, Kirsten Bagger, and Eva Brandt; Alison Willard and Richard Blyth from the British Telecom User Centred Design Group, Martlesham; Steve Bailey from Octo Design, Newcastle; Lynne Coventry and Michael McNamara from NCR in Scotland.

I would like to thank Mike Stott of Umeå University, Sweden and Alastair MacDonald, Glasgow School of Art for allowing me to present the design tools to their students and Gilbert Cockton, Martin Lavery of Sunderland University and Thomas Green of Leeds University for their input into the development and adaptation of cognitive dimensions.

The majority of the empirical research was carried out while teaching in the Institute of Design at Teesside University. I would like to thank Paul van Schaik for his incisive support during the early studies; Ben Norris for undertaking the prototyping work for the novel interfaces, his skills proved invaluable to produce such high quality and robust



prototypes; Katie Stockton, Andy Pearson and Claire Cooke for providing research support; and finally Steve Baskerville and Juliet Walters who allowed me to have the time and space for many of the studies to be carried out.

Most of the writing-up of this thesis was carried out at Huddersfield University and I would like to thank Janet Finlay, Robert Ward and in particular Phil Marsden for their criticism and support.

I would like to give a big thank you to Mark Porter for supervising this research and Phil Roberts for his direction.

I would also like to thank my wife, Sheralyn, who has unstintingly provided encouragement and support over the last seven years and, on occasions, took the children away for weekends to give me time.

Finally, I would like to thank my five-year old son, James, for being considerate and always asking, 'Is this your PhD?' before drawing on any scrap paper.