Physicality in the Design and Development of Computer Embedded Products

Joanna Hare

BSc (Hons) Industrial Design and Engineering

Director of Studies	Prof. Steve Gill Director of Research, Cardiff School of Art & Design, Cardiff Metropolitan University, Cardiff
Supervisor	Dr. Gareth Loudon Principal Lecturer, Cardiff School of Art & Design, Cardiff Metropolitan University, Cardiff
Supervisor	Prof. Alan Lewis Consultant to the Deputy Vice-Chancellor, University of Wales Institute, Cardiff

Submitted for the Degree of PhD by Research

Awarding body: The University of Wales

Date of Submission: 9th July 2014

This research was undertaken under the auspices of Cardiff Metropolitan University

Declaration

This work has not previously been accepted in substance for any degree and is not being currently submitted in candidature for any degree.

Signed………………………………………….. (candidate)

Date……………….

Statement 1

This work is the result of my own investigations, except where otherwise stated.

Other sources are acknowledged by footnotes and through explicit references. A full list of references is appended.

Signed………………………………………….. (candidate)

Date……………….

Statement 2

I hereby give consent for my work, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed………………………………………….. (candidate)

Date……………….

Abstract

Computers have become commonplace in our daily lives; they are embedded within so many products that they have largely become invisible. Furthermore, computers are used to aid in the design of those products and it is possible for the entire design process to be performed digitally. But humans are physical beings; evolved to have an innate understanding of the physical world. In contrast, this digital world is very new.

This research is an exploration of physicality in relation to the design and development of computer embedded products. Physicality is loosely defined for this thesis as the physical aspects or qualities of both an object and its interaction; this includes our physical bodies in relation to that object. The physical manifestations, or prototypes, used during the design of computer embedded products need to appear responsive to a user's action. These prototypes can be made interactive through embedding electronics within the prototype or 'faking' the interaction.

At the core of this research are two extensive studies for which a series of prototypes were created to answer the research question: **can a better understanding of physicality help in the creation of more effective low-fidelity physical interactive prototypes?**

These studies uncovered significant new knowledge into the role of physicality in the design of computer embedded products. Specifically, the notion of active and passive physicality is proposed.

Results suggest that, with a better understanding of active and passive physicality, designers can make more effective interactive prototypes for early stage user trials. Comparison of all the prototypes constructed revealed insights suggesting that the most effective prototypes balance both active and passive physicality equally. In addition, the notion of physicality can demonstrate why, in these studies; paper prototyping, screen-based prototypes and even Arduino prototypes produced unsatisfactory user data.

Acknowledgements

This thesis represents a seven year journey. I would like to give special thanks to my supervisors Prof. Steve Gill, Dr. Gareth Loudon and Prof. Alan Lewis who have provided invaluable guidance and support along this journey. Thank you.

I would also like to thank the Prof. Alan Dix, Prof. Steve Gill (again) and Dr. Devina Ramduny-Ellis (the Designing for Physicality (DEPtH) project team). The DEPtH team set the foundation for this thesis and provided many stimulating and thought-provoking discussions that really got me thinking about physicality and the role of research.

I would also like to thank both Dr. Alex Woolley and Dr. Ian Culverhouse who were part of the research team at Cardiff Metropolitan University (then UWIC) during the development of this research structure. Many interesting discussions during this time really helped to refine my thinking.

The National Centre for Product Design and Development Research (PDR) and Cardiff Metropolitan University have both provided valuable support along this journey. PDR has an active, lively and supportive research culture that has really inspired me, especially during the writing up phase. Through active involvement in the research student experience at both Cardiff Metropolitan University and the University of Wales I have gained valuable experience of the wider research environment. My thanks goes to those that have enabled this experience, specifically: Dr. Andy Walters, Jarred Evans (PDR); Jeff Alder and the Academic Associate Committee (Cardiff Metropolitan University); Prof. Robert Brown and the Research Degree Committee (University of Wales).

During this research I have conducted user trials with over 80 participants; I would like to thank you all for giving your time to participate in the user studies.

Thanks also to Christopher Wilkinson for proof-reading the thesis.

Finally, I would like to thank my family who have been very supportive and patient during my journey, especially my husband, Joe and my parents, Rod & Sandra. Thank you.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

Computing technology has transformed our world over the last 40 years. Since the first 'personal computer' in the 1980s, computers have become embedded into our everyday products. The integration of this 'invisible' computer adds an intangible aspect to our products with capabilities beyond the purely physical form. Technology has progressed to a level where users do not need to know they are using a computer, just that they are putting the washing on or making a phone call, thus moving the emphasis from technology-based products to user-centred products.

The design team of computer embedded products will usually comprise both Industrial (or Product) Designers and Human Computer Interaction (HCI) specialists. These products represent an interesting challenge for the design team; typically the physical product is designed by the Industrial Designer, while the software interface is the domain of the HCI specialist. However, to design a product that meets the needs of the user, the HCI specialist needs to recognise the difference between programming for a desktop computer and a computer embedded product (Bergman & Haitani, 2000, p. 2). And the Industrial Designer needs to understand more about programming and electronics in the recognition that they are no longer designing solely what the product looks like, how it feels in the hand and how it is made, but also "how it behaves" (Moggridge, 2007, p. xvi).

The design of any product is a complex process, if an inappropriate decision is made during the early stages, and this is not recognised until the latter stages of the process, major changes may need to be made to the design. Small modifications to an unsuitable design will not solve a major design issue, and the design will have to go back to the concept stage. Therefore, an opportunity exists at this early stage to aid the design team to work through ideas quickly and efficiently in order to find, and avoid, all the potential pitfalls that could be incurred by proceeding to the later stages too soon. However, computer embedded products are difficult to mock up in these early stages because of their combined physical form and digital interactions.

Low fidelity prototypes are often used during the early stages of the design process. The prototype should convey or explore what the design team needs, yet be so low-investment that it can be discarded once it has performed its function. The concept of fidelity is used to express this 'investment' in the prototype, with low fidelity being a lower 'resolution' of the model in relation to the final design intent. However, the literature review reveals that much of the research on the fidelity of interactive prototypes has focussed purely on the digital aspects of the prototype or, at best, it has given the physical prototype limited consideration.

This thesis is based on the proposition that concepts of ‘physicality’ could be used to address some of the effects of the increasing ‘digitality’ of the computer embedded product. The literature explores three interrelated themes of physicality: the meaning of physicality in relation to computer embedded products; physicality in relation to the design process of computer embedded products; and the physical manifestations of the design process (specifically prototypes).

The research presented in this thesis has its foundations in the Designing for Physicality (DEPtH) Project, funded through ‘Designing for the 21st Century’, a joint AHRC and EPSRC initiative. DEPtH was a collaborative research project between Lancaster University and Cardiff Metropolitan University (then known as the University of Wales Institute, Cardiff (UWIC)). The objective of the DEPtH project was to ‘*investigate the impact of physicality on product design - how humans experience, manipulate, react and reason about 'real' physical things, and how this understanding can inform the future design of innovative products’*.

Four workshops were organised in conjunction with the DEPtH project, three of these were part of the International Workshop on Physicality series and the fourth focused on physical fidelity. These workshops were; the ‘First International Workshop on Physicality’ (Ghazali, Ramduny-Ellis, Hornecker, & Dix, 2006), the ‘Second International Workshop on Physicality’ (Ramduny-Ellis, Dix, Hare, & Gill, 2007) held in conjunction with the British HCI Conference in 2007, ‘Physical Fidelity in Design: a shared exploration’ held in 2008 at Cardiff Metropolitan University and the ‘Third International Workshop on Physicality’ (Ramduny-Ellis, Dix, Hare, & Gill, 2009) held in conjunction with the British HCI Conference in 2009.

An account of the DEPtH project can be found in Chapter 11: ‘Design and Physicality – Towards an Understanding of Physicality in Design and Use’ (Dix, Gill, Ramduny-Ellis, & Hare, 2010) in the book: Designing for the 21st Century.

The research for this PhD has been conducted through the National Centre for Product Design and Development Research (PDR) at Cardiff Metropolitan University. PDR are a world renowned design and innovation consultancy and research centre with a wide variety of clients including designers of kitchen consumer products, baby products and medical devices.

With its foundations in the DEPtH project and the work of PDR, the research presented in this PhD thesis explicitly explores physicality in the context of designing complex, modern products; looking at how the physical impacts the way in which these products are designed. With an ever evolving stream of new technology and the latest gadgets, computer embedded products are a fast-paced topic but this research draws on the essence of the designed object and explores physicality at a fundamental level.

The approach taken to this research was as follows:

Literature review and development of the Research Question (Chapter 2)

A review of literature was undertaken along the three interrelated themes; theme 1: the meaning of physicality in relation to computer embedded products, theme 2: physicality in relation to the design process of computer embedded products, and theme 3: the physical manifestations of the design process (specifically prototypes).

'Contextual Studies', conducted as part of the DEPtH project, are presented within each theme. These Studies helped obtain a better understanding of physicality in relation to each theme.

As a result of the literature review and contextual studies the research question was refined to:

> ***Can a better understanding of physicality help in the creation of more effective low-fidelity physical interactive prototypes?***

Methodology (Chapter 3)

A methodological framework was created to address the research question with two studies at the core of the methodology. Each study focussed on a specific area of investigation; the first being fidelity and physicality, and the second, a framework for physicality. Both studies were based on the creation of a series of prototypes created to address the specific research aim of each study. These prototypes were analysed in terms of physicality before undergoing user trials. The results of Study One drove the direction of Study Two.

Study One (Chapter 4)

The first of the studies ('Conceptual Photo Management Product') investigated the relationship between fidelity and physicality. A series of prototypes was constructed using time constraints to drive the fidelity and the subsequent physicality of the prototypes. These prototypes were subjected to two phases of analysis; the first to assess their physicality and the second to assess their effectiveness through user trials. The hypothesis of this study was that there was a direct relationship between fidelity and physicality and that this relationship would affect the effectiveness of the prototype.

The results of this study indicated that there was a relationship between fidelity and physicality but that the relationship was more complicated than hypothesised. This led to a critical reflection of the study and a new hypothesis was proposed for Study Two, focusing on physicality rather than fidelity.

The new hypothesis was that a framework of better understanding physicality, and therefore controlling the levels of physicality in each prototype, would influence the effectiveness of that prototype.

This framework was based on the separation of physicality into active and passive physicality. Passive physicality is proposed to relate to the physical properties of the prototype and active physicality focuses on the physical experience of interacting with the prototype.

Study Two (Chapter 5)

The second study ('Media Player') investigated the direct effect of physicality by using the framework of active and passive physicality to drive the construction of the prototypes. As with Study One, these prototypes were subject to two phases of analysis; their resultant physicality and an assessment of their effectiveness through user trials. The results supported the hypothesis and indicted 'best practice' in the creation of effective prototypes for use in user trials.

Discussion and Conclusion (Chapter 6)

The Discussion Section of this thesis reflects on the results of all the user studies, including the 'Home Phone' of the contextual studies. In total, eleven prototypes were compared in relation to the hypothesis of a framework for active and passive physicality used in Study Two. These eleven prototypes were used to further probe the guidelines of 'best practice' indicted by the results of Study Two.

The thesis concludes with the framework of active and passive physicality, where:

> **Passive physicality is the perceived affordance mainly based on the visual appearance and tangibility of the prototype.**
>
> **Active physicality is the perceptible experience of interacting with the prototype.**

The framework proposes that both active and passive physicality should be considered on a scale of low to high and that prototypes which fall below certain levels of either active or passive physicality in relation to the design intent are least effective and prototypes that balance active and passive physicality equally are the most effective.

Chapter 2. Literature Review

2.1 Introduction

Computer embedded products have become commonplace in our daily lives. As users, we no longer need to know that we are interacting with a computer. Donald Norman's book 'The Invisible Computer' (1999) covers this topic extensively, highlighting that this is not the first technology to 'go invisible'. Norman provides the example of the electric motor when it was introduced to the domestic market in 1918 (1999, p. 55). This first electric motor was expensive, and was therefore marketed with attachments for many applications such as a food mixer, hair dryer or vacuum cleaner; so that the user would not need to buy several different motors. Today, the motor has become almost invisible; it has been embedded within a variety of products such as food blenders, printers and DVD players. The user does not need to know they are even using one, let alone be able to service it or understand the technology behind its operation. The same can be said today for the computer embedded product, users do not need to know they are using a computer, just that they are putting the washing on or making a phone call, moving the emphasis from technology-based products to task-based products.

Computer embedded products are dedicated to the task. The way the user interacts with the product is specific to the task performed, whatever that task might be; pushing buttons on a car stereo, turning a dial on a washing machine, moving your body in a specific way for the Xbox Kinect or by shaking an iPod Shuffle.

The integration of computing power adds an intangible aspect with capabilities beyond the purely physical form. These products interpret our actions electronically for any output to occur; and this interpretation does not have to conform to our experience of the physical world. For example, a product can defy physical laws by switching on seemingly without any interaction. The addition of computing power can allow this to occur through multiple means including a simple timer delay or by remote access to the product through Bluetooth or Wi-Fi.

The evolution of products to the modern day computer embedded product from the first handcrafted product, has been broken down into four 'eras' by Frens (2006), these are:

1. **The era of handcrafted products** – these products relied on the skill of the operator for their functionality. How a product should be operated could be 'read' from its form. For example, a pocket knife can be made on an individual level and the blade 'affords' cutting.
2. **The era of mechanical products** – as technology progressed, mechanics were added to accomplish tasks that humans could not readily do. These are the first products that required some sort of interaction through rudimentary controls. Use could be determined because the controls are inseparably part of the product. For example, a hand drill has a handle that turns the drill bit through mechanical means and the functionality can be easily 'read'.
3. **The era of electrical products** – electrical components gradually supplemented and replaced mechanical components. Electrical components were smaller than mechanical systems resulting in more freedom in the design of the product. Electrical products can be operated differently to mechanical products; with switches and dials instead of cogs and levers, introducing a level of abstraction to be understood when operating the product. For example, a radio has a dial which is turned to change the volume; the meaning of the dial is added through labelling the dial.
4. **The era of electronic products** – this occurred when electrical components were replaced by electronic components controlled by embedded computers. Electronic components are smaller than electrical components meaning that miniaturisation could happen on a greater scale. Components can be built into almost any product extending their potential use far beyond the purely physical. Interaction occurs through the operation of an interface which is interpreted by the embedded computer adding further to the level of abstraction. For example, changing the volume on a mobile phone can be executed a number of ways; these tend to be either on screen or dedicated buttons, neither of which is determined by the electronic hardware actually controlling the volume.

Our physical world has 'material existence' and is therefore 'perceptible, especially through the senses, and subject to the laws of nature' (Merriam-Webster, n.d. (a)). As humans, we have been shaped by millennia to interact with this physical world and later 'handcrafted' products. Yet a computer embedded product does not have to conform to that tacit knowledge of our physical world. The addition of a computer introduces an intangible element that has removed the physical linkage between cause and effect.

The interface of a computer embedded product provides the link between user and a product's functionality (Frens, 2006). However, despite the 'interaction' with the interface being at the core of a computer embedded product, the term 'Interaction Design' is not often used in this thesis. Interaction Design, with its roots in HCI, is a discipline in its own right and it is viewed by many as an entirely digital medium (Lowgren, 2013). Interaction Design is often discussed in the context of web design or the digital aspects of design; this thesis focuses on the point at which the physical and digital meet, not the discipline of Interaction Design.

This chapter provides a review of literature in relation to three interrelated themes;

1. the meaning of physicality in relation to computer embedded products
2. physicality in the design process of computer embedded products
3. the physical manifestations of the design process (specifically prototypes).

In addition, 'contextual studies' are presented within each theme. These studies were conducted as part of the AHRC/EPSRC funded 'Designing for Physicality' (DEPtH) Project. DEPtH was a collaborative project between Prof. Alan Dix and Dr. Devina Ramduny-Ellis at Lancaster University and Prof. Steve Gill and myself (J. Hare) at Cardiff Metropolitan University (formerly UWIC).

All of the contextual studies were conducted outside this thesis. The research was conducted by all four members of the DEPtH research team named above. Each of the studies has been published and the full papers are included in the Appendix. A diagram of the relationship between the literature review and contextual studies is shown in Figure 1. The contextual studies are included because of the role they played in informing the development of the research question.

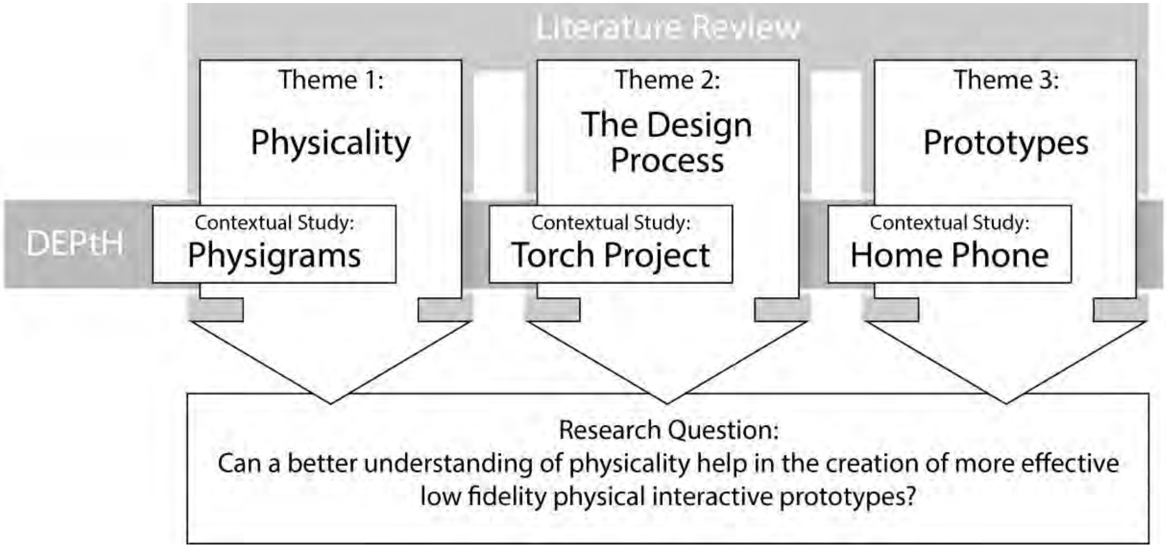


Figure 1: The relationship of the literature review and the DEPtH project

2.2 Theme 1: Physicality and Computer Embedded Products

Physicality is becoming a term that is increasingly used in the fields of Product Design and Human Computer Interaction (HCI). As a word it is easy to define. Physicality is the physical nature of something. Yet, when it comes to describing physicality in more depth, its specific meaning becomes ambiguous and it becomes a very difficult subject to understand. This is demonstrated in the proceedings of the Second International Workshop on Physicality in which Israel (2007) states *'At the 2006 Physicality workshop, a lot of contributions were concerned with the very meaning of physicality in the context of human-computer-interaction. But a consensus could not be found about what physicality is'.*

To build an understanding of physicality in relation to computer embedded devices three philosophical discussion areas are identified and explored;

1. humans as physical beings within our physical world (embodiment)
2. physical signifiers for interaction (affordances)
3. the point at which the digital and physical meet (interaction).

2.2.1 Physicality and Embodiment

Cognition, perception and the actions of humans are based on a complex interplay between our brains, our bodies and the world (Clark, 1998). Haugeland (1998) proposes that the mind is 'intimately *embodied* and intimately *embedded* in its world'.

Human *experience* of the world is central to the phenomenological branch of philosophy. Phenomenology uses the phenomena of existence to address philosophical questions of being and existing (ontology) and the study of knowledge (epistemology) (Dourish, 2001, p. 103). Perhaps the most relevant of phenomenological theorists is Merleau-Ponty (1945) who focused on the role of the body in perception and understanding (Dourish, 2001). Merleau-Ponty argued that perception of external reality comes through, and in relation to, a sense of the body. There is a relationship between the body and the world within which it exists; we are 'embodied' in the physical world. Embodiment refers to the understanding we have of the world though our experience of it (Dourish, 2001).

Embodiment recognises that we are influenced by our outside world (situated-ness) as much by our internal world. Situated-ness refers to the behaviour that results from interacting within a physical 'situated' environment. Our internal world is informed by our senses but also acknowledges that our emotions, desires and needs affect our perception of the world.

Humans have evolved for millennia to exist within the physical world; our bodies are a certain size and shape to enable effective interaction with that world. But it is not just the physical manifestation of our body within the world that informs our interaction with it. The way the world is *perceived* through our bodies' *senses* is also important; as is the interpretation of those senses. Our physical presence in the world is informed by vision, sound, touch, taste and smell; these senses also inform us of what we can do in, and to, the world; our 'potential for action' (Larssen, Robertson, & Edwards, 2006).

The sense of touch is perhaps the most directly physical. Consider a mug for example; the tactile form creates the fundamental physical experience because you can physically 'feel' it. Yet the visual sense can also inform the tactile sense, the mug could look heavy and easy to grasp or light and delicate. Reeves (2006) describes physicality as being 'shaped by the physical properties' of an object and explains that this is primarily informed by 'the tactile and visual senses'. Ghazali (2006) agrees with this in terms of 'pre-technology' where the physicality of an object was based *'solely upon its physicalness; the interpretation of what it is, or what we are supposed to do with it, depended heavily on its physical-bodily appearance'.* Pre-technology, in this case, relates to the time before technology during which humans evolved to fit the physical world. Because of our need to understand the physical world in order to survive, there exists a natural and implicit knowledge of that world. These implicit behaviours are naturally understood by all, a stone will fall, water will flow around an obstacle and fire will burn.

The use of technology to aid our everyday existence has progressed from handcrafted, Stone Age tools to help with survival, hunting and cooking, to the modern day dependence on technology which is omnipresent in our daily lives. It is the relatively recent advent of mechanical, electrical and electronic products that defy our implicit understanding of our physical world (Ghazali & Dix, 2005). A good explanation of this is that of Paul Watzlawick (1967, p. 29) as described by Israel (2007); if you kick a stone it will stop at a place that is predetermined by the amount of energy transferred plus the shape, weight and surface characteristics of the stone. Whereas, if it is a dog being kicked; it will react in a fundamentally different manner and the energy is not transferred 'logically', the energy might be translated into a 'fight or flight' response. This example demonstrates that 'intelligent' objects are less predictable and therefore require more attention and awareness than non-intelligent objects.

2.2.2 Physicality and Affordances

Affordances are the subject of a lot of debate in design literature (McGrenere & Ho, 2000) (Gaver W. , 1991) (Djajadiningrat, Overbeeke, & Wensveen, 2002). Donald Norman is credited with being the first to introduce the term to designers in his book *The Psychology of Everyday Things (POET)* (1988). Norman's version of affordances was appropriated from James J. Gibson's definition proposed in his seminal book *The Ecological Approach to Visual Perception* (1986). Norman used Gibson's concept of affordances and applied it to the designed world in an effort to increase awareness of how people interact with everyday things. However, the two definitions differ and despite Norman's version being used widely, there are ambiguities that have led to widely differing uses of the term.

Gibson was a perceptual psychologist and invented the term to convey all "action possibilities" available between an 'actor' and their environment. These possibilities are independent of the individual's ability to recognize them, but always in relation to the actor and therefore dependent on their capabilities. One of the examples given by Gibson (1986) is that of water; to a human, water does not afford walking on, but the surface tension offered by water means it does afford walking on by a pond skater.

McGrenere and Ho (2000) highlight three fundamental properties of an affordance as defined by Gibson:

1. An affordance exists relative to the action capabilities of a particular actor.
2. The existence of an affordance is independent of the actor's ability to perceive it.
3. An affordance does not change as the needs and goals of the actor change.

This definition disregards the knowledge and expectations of the actor; therefore Gibson focuses only on the *action capabilities* of the actor. Norman, on the other hand, proposes that the perceptual and mental capabilities of the actor affect the affordance, therefore introducing the distinction of actual versus *perceived* affordances. An example of this is the American light switch which has a twist switch; to both a British and American user, experience dictates that this should be twisted for the light to come on. When it comes to turning the light off the experience of British lights dictates that the reverse action should be performed, however, this is not the case, in fact the switch should be twisted further causing many British users to get this wrong. Therefore the perceived affordance of the twist switch differs based on the actor performing the action, despite the actual affordance remaining unchanged.

Since the first introduction of affordances in POET, Norman has issued an article on his website (2004) stating that his use of the term affordances should be regarded as perceived affordances, because:

"(As designers) we care much more about what the user perceives than what is actually true. What the designer cares about is whether the user perceives that some action is possible"

In fact, he now refers to perceived affordances as **signifiers** (Norman D. , 2013). Norman recognised that the term 'affordance' in design had come to mean something perceptible. Perceived affordances and signifiers are methods for communicating potential for action.

Much of the ambiguity has arisen from interpretation of visual perception; does an affordance exist if the actor is not aware of it? When Norman applied Gibson's idea to design; he divided the idea of affordances into those of *real* and *perceived* affordances. Whilst real affordances are what the user could actually do with the device, meaningful or not, perceived affordances tell the user '*what actions can be performed on an object and, to some extent, how to do them*'. To explain, when a user perceives the affordance, the visual information also gives some indication of how to act out the affordance. For example, thin vertical door handles afford pulling while flat horizontal plates afford pushing (Gaver W. , 1991). When grasping a vertical bar, the hand and arm are in a configuration from which it is easy to pull; when contacting a flat plate, pushing is easier.

Norman's use of the term affordances is based on perception, there is either perceptual information available to the user or there is not. Gibson's use of the term is based on affordances existing (or not) independent of the actors being aware of it. Gaver (1991) has clarified this further creating a visual diagram by separating affordances from the perceptual information available about them, as shown in Figure 2 below, thus dividing affordances into three categories: perceptible, hidden, and false (the final quadrant being correct rejection of a non-affordance).

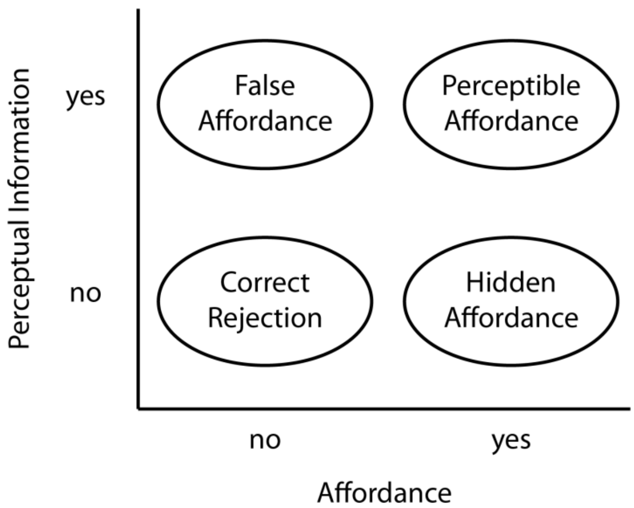


Figure 2: Separating affordances from the information available about them allows the distinction among correct rejections and perceived, hidden and false affordances (redrawn from Gaver, 1991).

A 'correct rejection' and 'perceptible' affordance are the two desired outcomes in designed objects. In the case of a correct rejection, there are no 'action possibilities' and there is no perceptual information suggesting an action, therefore the actor does not attempt any action. For a 'perceptible' affordance, both the affordance and the visual information exist for the actor to perform a successful interaction. For a 'false' affordance, visual information exists that the actor interprets as actionable, yet none exists, for example a door that the user tries to go through but is found to be locked. Finally, a 'hidden' affordance is one that cannot be perceived, for instance a user has to have knowledge that a hidden door exists, independent of what their senses perceive.

Hartson (2003) applied this to digital interaction design introducing four different kinds of affordances; cognitive, physical, sensory and functional affordances. A cognitive affordance enables the thinking or knowing about something; a physical affordance enables physically doing something; a sensory affordance uses the physical senses, and finally a functional affordance aids the user in doing something. A sensory affordance is often an attribute of cognitive and physical affordances; a functional affordance should always be part of a physical affordance. However, the proposal by Hartson specifically addresses on-screen, digital 'affordances'. The research presented in this thesis is interested in affordances of the physical world therefore Norman's definition of perceived affordances will be used, or what he is now calling signifiers.

2.2.3 Physicality and Interaction

Physicality is not just a property of objects; it is also an effect of interaction argues Büsher (2006). No object can be used without some sort of physical interaction and therefore physicality will have an effect. In this research, interaction is regarded as human action at the point at which the physical and the digital meet. In order to use a computer embedded device, interaction with it needs to take place in some way, from navigating a car to washing clothes in a washing machine or using a mobile phone to make a phone call.

Svanaes (2013) defines the interactivity of a product as 'the way in which it responds to actions by a user'. Dourish (2001) considers interaction not only as what is being done, but also how it is being done; interaction is the means by which work is accomplished, dynamically and in context.

Ghazali and Dix (2005) propose three inherent properties of physical objects in comparison to software; they refer to these as 'principles of physicality'. Two rely on interaction; these are 'directness of effort' and 'locality of effect'. The third informs the user of potential actions, its 'visibility of state'.

- **Directness of effort**, a small amount of effort results in a small effect whereas a large amount produces a large effect (for example, throwing a ball).
- **Locality of effect,** the effort applied to a physical object will affect it in that location, if you throw a ball now, it will move now, not in five minutes time.

- **Visibility of state,** as a result of the above a physical object will operate within the laws of physics that we have grown to accept, thus demonstrating their 'state'.

Part of the complexity of computer systems, continue Ghazali and Dix, is that they violate these principles of physicality.

Ozenc *et al.* (2010) characterise controls, which they define as supporting communication between people and computational systems, as having four aspects; **affordance** (communication of action possibilities), **feedforward** (communication of the outcome, before any action is taken), **expression** (ability of the user to express their intention to the system), and **feedback** (communication that the system has recognised the user's action).

Larssen *et al.* (2006) presented five aspects that the designer should consider in relation to the *feel dimension* of technology interactions. They isolate the feel dimension as how we use our 'inner' sense and motor skills when incorporating a tool into our bodily space so that it becomes an extension of our bodies. The five aspects they propose are body-thing dialogue, potential for action, within-reach, out-of-reach and movement expression. Body-thing dialogue is described as the dialogue between our perception and the 'thing'. It is an interplay between our bodies and the world available to us, where our bodies are engaged in a dialogue with the 'thing', allowing and enabling the second of the proposed aspects: potential for action. The ways in which interactions occur are proposed to be either 'within reach' or 'out of reach', which describes tangible versus intangible interactions and the resultant feedback loop. 'Movement expressions' refers to how movements are executed to create the dialogue between the body and the thing. To interpret this in relation to a mobile phone:

- Body-thing dialogue - the mobile phone is grasped and bought into our bodily space
- Potential for action – we have hands with fingers which can potentially be used to interact with the product (or indeed perform an unintended action such as throwing it)
- Within-reach – we can perform an action directly on the product in order to achieve something
- Out-of-reach – for example, lack of mobile phone signal or inability to operate an interface through arthritis, mean actions do not have the desired effect

- Movement expression – the way in which actions are performed can be expressed at an individual level, for example the same action can be performed whether the mobile phone is used in the left hand, right hand or both.

In all cases, interaction is seen as being greater than simply the point at which the interaction occurs. Interaction is concerned with what is happening before, during and after the interaction even on purely mechanical devices. As an added complication for computer embedded products, the natural 'cause and effect' properties can be broken causing the product to become difficult to learn how to use.

2.2.4 Summary

Physicality is concerned with the physical aspects or qualities of both an object and interaction; this includes our physical bodies in relation to that object.

Physicality is central to our experience of all devices and computer embedded products are no exception, from how we exist in our bodies within the physical world, through how we perceive interactions with the physical world, to the point at which we interact with that physical world.

The Contextual Study presented in the next section develops the understanding of physicality in relation to computer embedded products by regarding the physical qualities separated from their digital actions.

2.2.5 Contextual study –'Physigrams'

The full paper can be found in Appendix 1:

> Dix, Ghazahil, Gill, Hare, & Ramduny-Ellis (2009), **Physigrams: modelling products for natural interaction**. In Formal Aspects of Computing, Volume 21, Number 6, December, 2009.

This study explored a very specific area of the meaning of physicality in relation to computer embedded products; namely, how the physical aspects can be formally represented for computer embedded products, the resultant diagrams are called 'Physigrams'.

2.2.5.1 *Development of the Physigram Notation*

This research set out to explore the formal representation of the interactive behaviour of physical products. The idea central to this investigation was to consider the product 'unplugged' and entirely separate from any digital or other external functionality. For example, if a mobile phone has run out of battery, or a light switch is unscrewed from the wall; they will still both have physical interactive behaviours: the phone buttons can be pressed and the light switch can be operated, even though there is no resulting effect on the phone or light. By separating the physical from the digital, a better understanding can be developed of the relationship between the two.

Feedback is a critical factor of interaction, and when the physical product and the digital effects are considered separately, the different ways in which users get feedback from their actions can be determined. For example, if a mouse button is pressed, the user both feels the button go down and sees something happen on the computer screen.

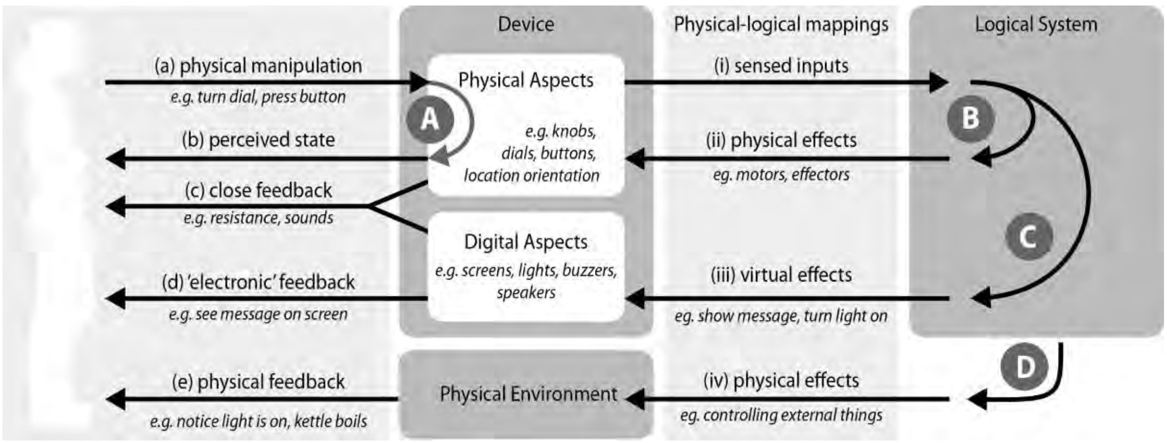


Figure 3: Feedback loops

Dix uses Figure 3 to show some of these feedback loops for a non-separated product. Interactions typically start with some physical action (a). This could include making sounds, but here the focus is on physical actions such as turning a dial, pressing a button or controlling a mouse. In many cases this physical action will have an effect on the product: the mouse button goes down, or the dial rotates and this gives rise to the most direct physical feedback loop (A) where you feel the movement (c) or see the effect on the physical product (b), or a combination of the two.

Dix describes that the user's physical actions must be detected by the product in order for there to be any effect to the digital system (i). For example, a key press causes an electrical connection detected by the keyboard controller. This may give rise to a very immediate feedback associated with the product; for example, a simulated key click or an indicator light on an on/off switch (ii). In some cases this immediate loop (B) may be indistinguishable from actual physical feedback from the product (e.g. force feedback as in the BMW iDrive); in other cases, such as the on/off indicator light, there is no physical effect but an effect which appears part of the product due to its proximity and immediacy.

The sensed input (i) will also cause internal effects on the logical system, changing its 'internal' state; for a GUI interface this may be changed text, for an MP3 player a new track or increased volume. This change to the logical state can cause a virtual effect (iii) on a visual or audible display; for example a screen showing the track number (iii). A feedback loop which is meaningful to the user is created (C) when the user perceives these changes (d). In direct manipulation systems the intention is to make this loop so rapid that it feels like a physical action on the virtual objects.

Finally, some systems affect the physical environment in more ways than changing screen content. For example, a washing machine starts to fill with water, or a room light goes on. In addition there may be unintended physical feedback, for example, a disk starting up. These physical effects (iv) may then be perceived by the user (e) giving additional feedback and creating a fourth feedback loop (D).

To focus on the purely physical aspects of interaction, consider the product 'unplugged' or without power. In this situation, the action of the user cannot be sensed by the product, and the 'internal' state will not be affected by the physical action of the user. The user will not perceive any electronic feedback, such as a menu item changing, or any physical feedback resulting from electronic intervention, such as a washing machine filling up (feedback loops C and D).

The paper goes on to propose augmentation of basic state transition networks (STNs) to enable representation of physical properties in addition to the state of the product. The resultant diagrams are called Physigrams.

2.2.5.2 *The Study and Outcomes*

The Physigram notation was trialled with designers who applied it to four prototypes which explored interaction with the same underlying application; this choice enabled comparison of superficially similar, but subtly different products through differences in the Physigrams.

The resultant Physigrams were analysed by the developers to determine how the designers had interpreted the use of the notation. The Physigrams did indeed depict the subtle differences in the physicality of the prototypes however, there was some ambiguity regarding viewing the prototype 'unplugged'. The designers had used the Physigrams to specify the intended design not the limitations of the prototype. Therefore it was proposed that there are actually three levels when considering a device 'unplugged':

1. The physical prototype fully connected with its underlying application
2. The prototype with its internal electronics, but unplugged from its application
3. The purely physical aspects of the prototype

The Physigrams were originally intended for level 3, but the designers had used them for level 2, both are important. This may mean that for certain prototypes we should consider drawing both level 2 and level 3 diagrams or perhaps annotating a single diagram to make clear which elements are level 2. Level 2 Physigrams embody knowledge that is available to the designers, but may not be apparent to users.

2.2.6 Conclusion

This theme explored the meaning of physicality in relation to computer embedded products. The philosophical discussion areas identify the importance of the individual in the perception of physicality.

Embodiment recognises the influences of the outside world (situated-ness) in relation to the internal world (emotions, desires, needs). In addition, an individual's perception of that world is informed by their senses. These senses provide information of what can be done in, and to, the world; the individual's 'potential for action'.

Embodiment and perception can be affected by the use of tools and technology. The individual's potential for action in the world is based on the perception of affordances to indicate what can be done within the physical world.

Finally, there is the point at which interaction with the world occurs; this interaction is seen as being greater than simply the point at which the action occurs. Interaction is concerned with what is happening before, during and after the interaction even on purely mechanical devices. For computer embedded products, natural ‘cause and effect’ properties can be broken, causing the product to become difficult to use.

The Contextual Study (‘Physigrams’) explored a very specific area of the meaning of physicality in relation to computer embedded products. Although not the primary intention of the study, this research provided insights on physicality through the consideration of the product ‘unplugged’. This enables reflection on the purely physical aspects of interaction with computer embedded devices which proved vital in the later stages of this thesis.

This theme allowed exploration of the meaning of physicality in relation to computer embedded products. The research question will seek to further investigate the importance of addressing the affordances of the prototype and the potential for considering the prototype ‘unplugged’.

2.3 Theme 2: The Design of Computer Embedded Products

The two disciplines of Industrial Design and Human Computer Interaction will be introduced and discussed in this section, giving a brief history to provide the context of the design of computer embedded products.

The shift in focus from technology-driven to task-driven, and therefore user-centred, products is then discussed. Subsequently, the typical product design process is introduced which then focusses on the early stages of that process. Tools and techniques specific to the user centred design process are then discussed, and the challenges specific to the computer embedded products are identified.

Finally, there are a number of opportunities for physical manifestations of the intended design during the design process. Römer *et al*. (2001) use the term external representations of the design process to convey the ‘outcomes’ of the design process both in a physical and digital format. These are the points at which physicality can be related to the design of computer embedded products.

2.3.1 Industrial Design (ID)

Humans have been ‘designing’ for millennia; the consideration of how an artefact is realised is demonstrated in the earliest cave paintings and flint tools. Design has roots in the long tradition of craftsmanship in which the conception and realization of an artefact is conducted by one person, or a small group of people, shaping materials by hand and eye (Heskett, 1980, p. 7).

Industrial Design is a derivative of design which Heskett defines as the "*process of creation, invention and definition separated from the means of production*".

In the mid-20th century the role of 'design' became to translate the output of the companies' research and development department into 'a form accessible and acceptable to the public' (Heskett, 1980, p. 142). However, this approach allowed design to become separated from the rest of the product development team; a product was first 'engineered' and later 'styled' by the designer. This is especially apparent with the advent of electronic products which have a level of abstraction in their interface. Interaction with the interface is not comparable to the design of the form or the way in which the product is manufactured, both of which Industrial Designers are traditionally skilled at undertaking. Therefore, the design of the form and physical aspects were undertaken by the Industrial Designer while the development of the electronic interaction and interface was undertaken by the Human Computer Interaction (HCI) specialist (Frens, 2006).

As technology progressed, products became feature-intensive, adding more and more functionality to already complex systems. The tipping point came with the realisation that products are ineffective if the user cannot interact with them. The flashing clock on an oven or VCR are well known demonstrations that the inoperability of the product has a direct effect on the users' experience of that product. In both cases the user cannot take full advantage of the product; the oven and VCR cannot be set to turn on at a predefined time without knowledge of how to do so.

Electronic technology had reached a point where it could easily give the performance demanded by the 'average' user, but the user could not access it, and so the emphasis shifted from the technology back to the person using it. Interactions should be 'designed' as part of the product as a whole, and the Industrial Designer is best placed to undertake this design as a 'whole'.

In the present day, there is some dispute concerning the difference between a Product and Industrial Designer. Brunel University differentiate Industrial Design by *"including technical content whilst maintaining a creative and practical approach [of Product*

Design]" (Brunel University London, 2014). Yet Ulrich and Eppinger (2012, p. 209) point out that Industrial Design has historically been used to 'style' a product after its technical features were determined. For the sake of this research the terms Industrial Design and Product Design are considered interchangeable. The modern day Industrial Designer tends to base their way of working and thinking on internal knowledge, on changing the world and on a human perspective (Bartneck & Rauterberg, 2007). They are interested in turning human values into requirements and, subsequently, solutions. Designs are intended to be used, therefore the context of use must be understood to understand and evaluate them, likewise, the aesthetic effect of designs will be determined by the use of the product and the attitude of the users in relation to the product (Heskett, 1980, p. 174). A good Industrial Designer considers the design of the entire product. Whilst not necessarily being a specialist in all aspects, such as manufacturing principles, they know enough about all the required aspects, providing the opportunity to combine them and create something entirely new (Overbeeke & Hummels, 2013).

2.3.2 Human Computer Interaction (HCI)

In contrast with Industrial Design, the field of Human Computer Interaction (HCI) is very new. It was only with the advent of the microchip that HCI could emerge in the early 1980s. Initially this was as a specialist area within computer science that embraces cognitive science and human factors engineering. Before then only technology professionals and dedicated hobbyists interacted with computers. The 1980s marked the emergence of the Personal Computer (PC), complete with personal software (productivity applications, such as text editors, spreadsheets and interactive computer games) and personal computer platforms (operating systems, programming languages, and hardware). This made computers available to the world, making everyone within it a potential user. However, the 'average user' encountered many difficulties not experienced by the 'trained user' (Cockton, 2013). Some form of development was needed in terms of the 'usability' of the Personal Computer for it to become mainstream.

In its early days, HCI was primarily concerned with usability; one of the seminal books of this time is 'Usability Engineering' by Jacob Nielsen (1993). Usability became important for the design of any interactive software which was not intended to be used by a trained operator. Usability testing is a form of user testing that looks at how a user undertakes specific tasks and quantifies the performance of the product. Through systematic appraisal of the interface, usability engineering intends to enable the people who use the product to be able to do so quickly and easily in order to accomplish their own tasks (Dumas & Redish, 1999).

But early HCI development "took on the trappings of the traditional computer model" by looking at the world in terms of plans, procedures, tasks and goals (Dourish, 2001). A number of concurrent developments enabled the progression of computing to be based on understanding, and better empowering users (Carroll, 2013). These developments included:

- Cognitive science provided a framework through which humans could be understood by engineers. This incorporated cognitive psychology, artificial intelligence, linguistics, and the philosophy of the mind.
- Human factors engineering provided empirical and task-analytic techniques for evaluating human-system interactions.
- The advancement of technology enabled the emergence of computer graphics and information retrieval and allowed software engineers to focus on non-functional requirements of usability, such as the graphical display of information.

Until relatively recently, HCI has been confined to the desktop as personal computers were large and unwieldy. Interaction was based on the standard keyboard and mouse input products. However, the 'era of electronic products' has enabled miniaturisation of computers, driving HCI 'beyond the desktop' and onto specialist products such as mobile phones. Interaction is no longer confined to a mouse and keyboard and can be through a variety of different means. The transition has not been entirely successful because many of the founding principles of HCI were established when HCI was confined to the desktop. The HCI discipline has begun to recognise this and established the user experience as a fundamental principle. One of the effects of this recognition is the integration of the Industrial Designer into the design team.

2.3.3 User Centred Design (UCD)

User centred design (UCD) is a design philosophy that puts the user at the centre of the design of a system, service or product. This is being championed by both the Industrial Design and HCI communities. The emphasis on the needs of the user, as opposed to the product, is a fundamental change to the approach of design (Boztepe, 2007). The way in which the user is involved can range from inclusion at specific points in the design process, for example specification requirements and user trials, to inclusion as part of the design team for the entire process, known as Participatory Design (Ehn, 1988). UCD is both a broad philosophy and a variety of methods (Abras, Maloney-Krichmar, & Preece, 2004). Section 2.3.6 starting on page 28 discusses a variety of ways in which the user could be involved during the design process.

Designers are rarely typical users (Thimbleby, 1991) (Norman D. , 1988). For example, when designing a medical device, the designer can try and understand the context of use, but they will never be the primary intended user. Even for a product that the designer might be the end user of, such as a mobile phone, the skills and interests of the designer will rarely be 'typical', with a likely skew towards a 'lead user' (Von Hippel, 1986) or 'early adopter' (Rogers E. M., 2010).

The UCD process recognises that the people who will be using the products (or services) have a better understanding of what their needs, goals and preferences are. It is the designers' role to elicit this information by involving users at every stage of the design process (Saffer, 2010). Rudd *et al*. (1996) agree, saying that users do not know how to articulate their requirements (since much of their relevant knowledge is tacit and not accessible to conscious thought), and verbalising their requirements is not objective (Blackler, 2009). One of the common arguments against UCD is that users do not know what they want, and therefore it is difficult for them to inform the design process of their needs. Steve Jobs and Henry Ford are the commonly quoted opponents for the user centred design process. Steve Jobs stated that *"people don't know what they want until you show it to them"* during an interview with BusinessWeek (Jobs, 1998). Henry Ford supposedly stated that if he had asked people what they wanted in the early days of the Ford Motor Company then they would have said a faster horse (interestingly it seems that he did not actually say this (Vlaskovits, 2011) but it illustrates the argument well). In response, a UCD approach does not advocate doing exactly what a user says; instead, the UCD practitioner should ask why the user is saying or doing something, and interpret it into something new. Steve Jobs and Apple understand the needs of their user at a fundamental level and base new designs on those needs, not necessarily on what the user says they want (Bowles, 2011). In the case of Henry Ford, if he had asked why people were saying they wanted a faster horse, he would have received valuable insight into their desire to travel somewhere faster.

UCD is undertaken by professionals who use a variety of methods to gain an insight into why a user says or does something, and indeed some of the most interesting insights can arise when what a user actually does is different to what they say they do. UCD is about using a variety of user research methods in order to gain an understanding of user needs that can be interpreted into design insights to inform design.

2.3.4 The Product Design Process (PDP)

The product design process is a complex process by which an idea moves from concept to market. There are several books covering the topic in great detail (Ulrich & Eppinger, 2012), (Baxter, 1995), (Wright, 1998). A simplistic overview of the PDP is proposed by Ulrich and Eppinger (2012), this is shown in Figure 4. The output of the **planning** phase is the project mission statement which specifies the target market for the product, business goals, key assumptions and constraints. During **concept development** the needs of the target market is identified and product concepts are generated and evaluated.

System-level design defines the product architecture which includes identification of subsystems and components. **Detail design** specifies the geometry, materials and tolerances of the entire product. During **testing and refinement** prototypes are constructed which enable evaluation of the design for various reasons including assembly, tolerance checking and user testing. Finally, during **production ramp-up** the product is constructed using the intended production system enabling training of manufacturing staff and a final check of the product before mass-production.

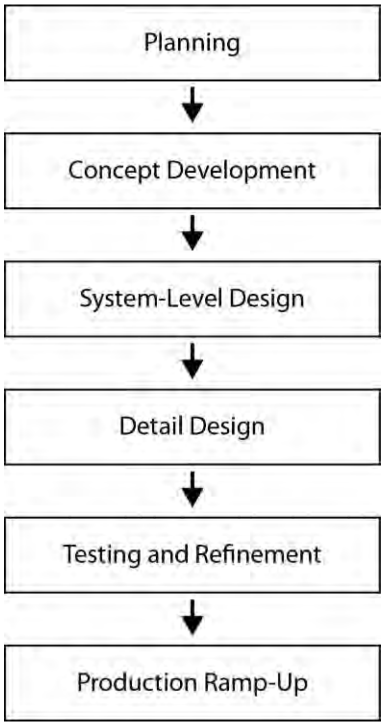


Figure 4: A simple overview of the product design process (redrawn from Ulrich and Eppinger (2012, p. 14))

A user centred approach advocates considering user needs from the earliest stages in the process throughout each stage. New design philosophies advocate an iterative design process where designs are trialled continuously by building prototypes, learning from them to inform the design process and move the design forward (Buxton, 2007). The cycle of design, prototype and evaluate is used repeatedly throughout the iterative process enabling a variety of opportunities at which to involve potential users. A generalised iterative design process is shown in Figure 5 below (redrawn from Interaction Design (Rogers, Sharp, & Preece, 2011, p. 332)).

Throughout the design process the designer will interact with materials, and there is a long tradition of designers manipulating physical materials, shaping them by hand and eye. Engagement with, and the manipulation of, materials is an intrinsic part of the design process as exemplified by the 'era of the handcrafted products' (Frens, 2006). This material-centric approach has shaped our modern understanding of design (Pevsner, 2005) (Raizman, 2010). Potter (1969) describes the activities of the designer as being driven by the manipulation of materials in workshops in which they 'get their hands dirty'. And when design became industrialised, clay, wood, metal and plastic were utilised because of their 'hands on' qualities (Heskett, 1980).

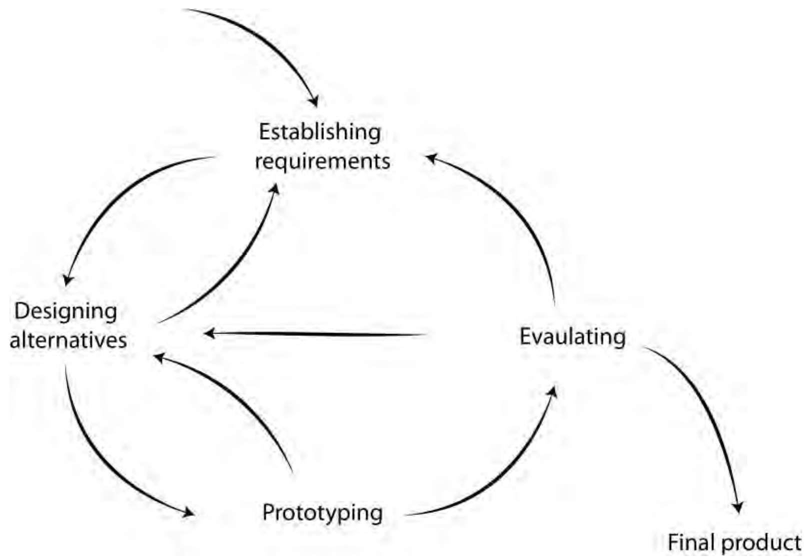


Figure 5: An iterative design process (redrawn from Rogers *et al*. (2011))

Yet when computers, and the associated software, began to be used in the design process, they bought virtuality, a stage that cannot be physically 'felt' or explored. Sketches and clay models gave way to Computer Aided Design (CAD) and virtual models. The connection with physical properties in the physical world was broken. This affected both the design process and the products being designed.

2.3.5 The Early Stages of the Design Process

"Because the investment in the product is low, the front end is the one time in the product pipeline when one can actually afford to play, explore, learn and really try and gain a deep understanding of the undertaking." (Buxton, 2007, p. 139)

The 'fuzzy front end' (FFE) is a term that developed in the 1980's to represent "the period between when an opportunity is first considered and when an idea is judged ready for development" (Kim & Wilemon, 2002). Reinertson (1999) relates the development of new products to a betting process, where the FFE is used to work out the associated risk compared to the potential return.

Baxter (1995) has constructed a diagram (Figure 6) showing when expenditure is committed verses the cost (including time allocation) up until that time. This demonstrates that relatively little expense is incurred during those early stages but that most of the decisions will have been made that have an impact to the cost of the whole project.

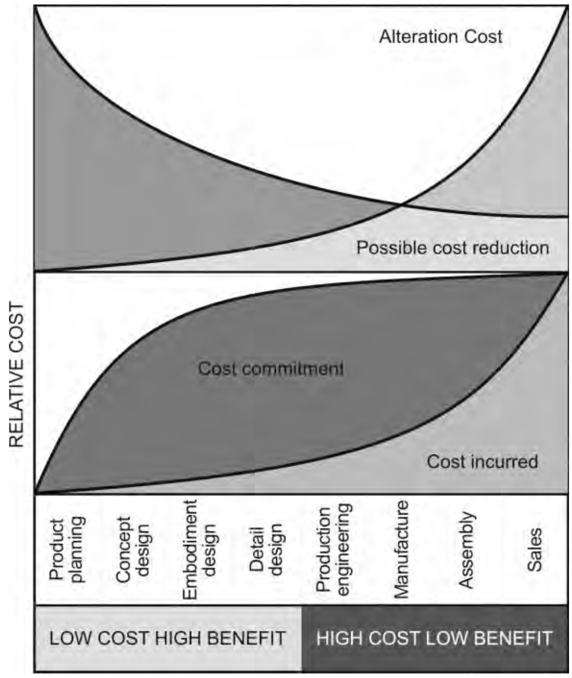


Figure 6: Costs and Benefits of the design stages (Baxter, 2002)

If an inappropriate decision is made during the early stages, and this is not recognised until the latter stages of the process, major changes may need to be made to the design. Small modifications to an unsuitable design will not solve a major design issue, and the design will have to go back to the concept stage. It is for this reason that this thesis focuses primarily on the initial stages of the design process, where the designer needs to work through ideas quickly and efficiently in order to find and avoid all the potential pitfalls that could be incurred by proceeding to the later stages too soon.

2.3.6 Involving the User

A review of literature reveals a host of tools and techniques that can be utilised during the design process for a variety of needs. Figure 7 shows some of the methods identified in the literature, and attempts to classify them according to their use within the design process; **generative research methods** gather data directly from the user, **interpretation methods** interpret that data into a form that can be discussed within the project team, **specification methods** capture the data in a more formal document, **ideation techniques** realise a concept and finally, **user trials** enable conversations with the user around a realised concept.

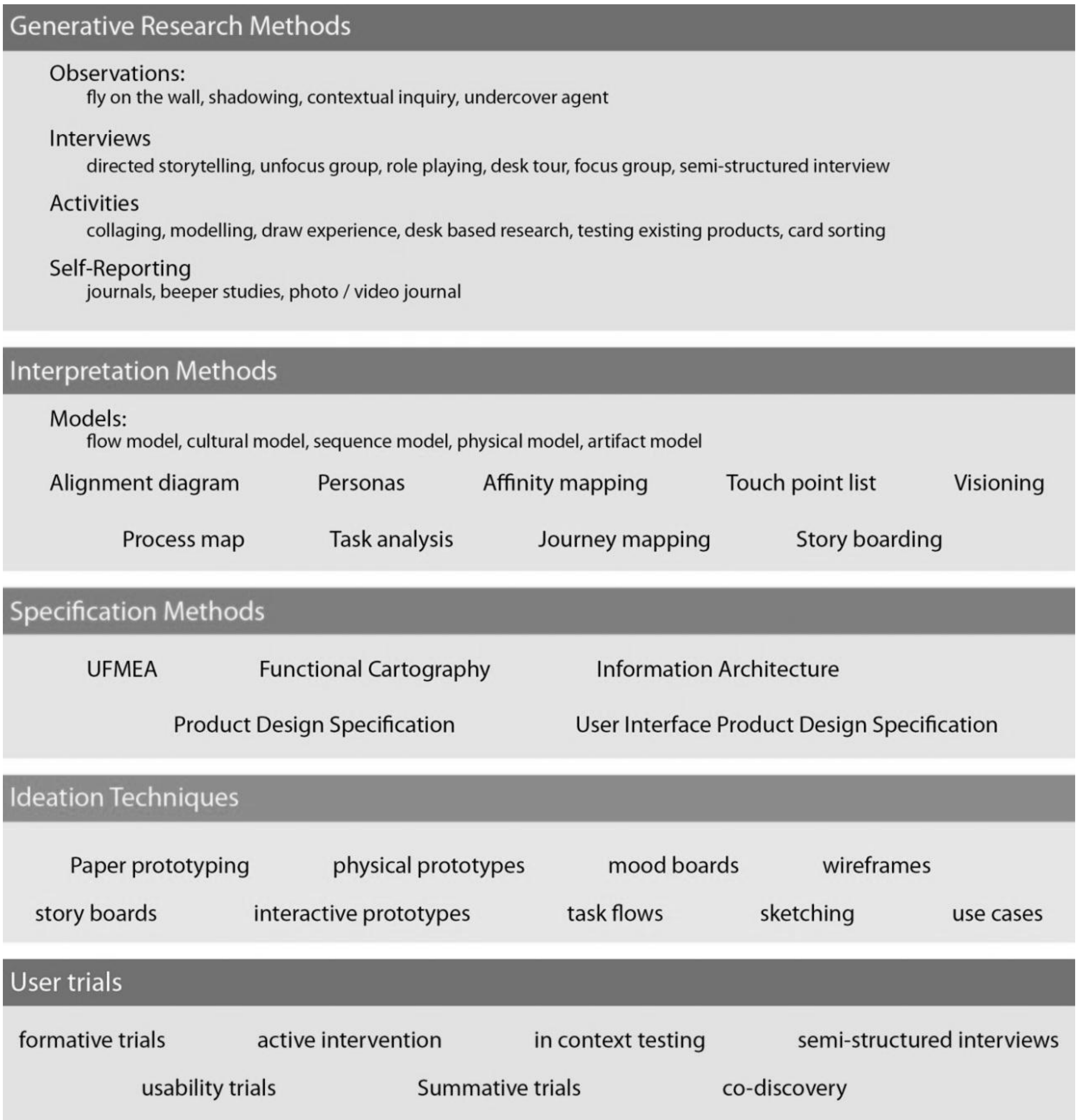


Figure 7: An overview of UCD methods identified in the literature

The generative research methods have been divided into four categories (Saffer, 2010); observational methods are used to observe what people do in a 'conscientious' manner, interviews involve talking to people, activities enable the designer to engage with an artefact alongside the user and self-reporting techniques are for the user to capture their own thoughts and activities without a researcher present.

Figure 8 (below) shows a simplistic overview of a UCD approach to an interaction design project typical of those undertaken by PDR. This Figure shows that data is generally collected through a combination of the generative research methods. For instance, a project might start with desk-based research looking at blogs, videos and product websites in order to give an overview of the product area and aid in potential user identification. From this the ethnographic observation technique, contextual inquiry can be undertaken to examine users going about their normal activities within the context of their work environment (Wixon, Holtzblatt, & Knox, 1990). Any questions arising from the contextual inquiry can be addressed in a semi-structured interview. Finally, existing and competing products can be assessed through heuristic evaluation or a user trial. Heuristic evaluation is a systematic inspection of a User Interface from a usability perspective (Nielsen & Molich, 1990).

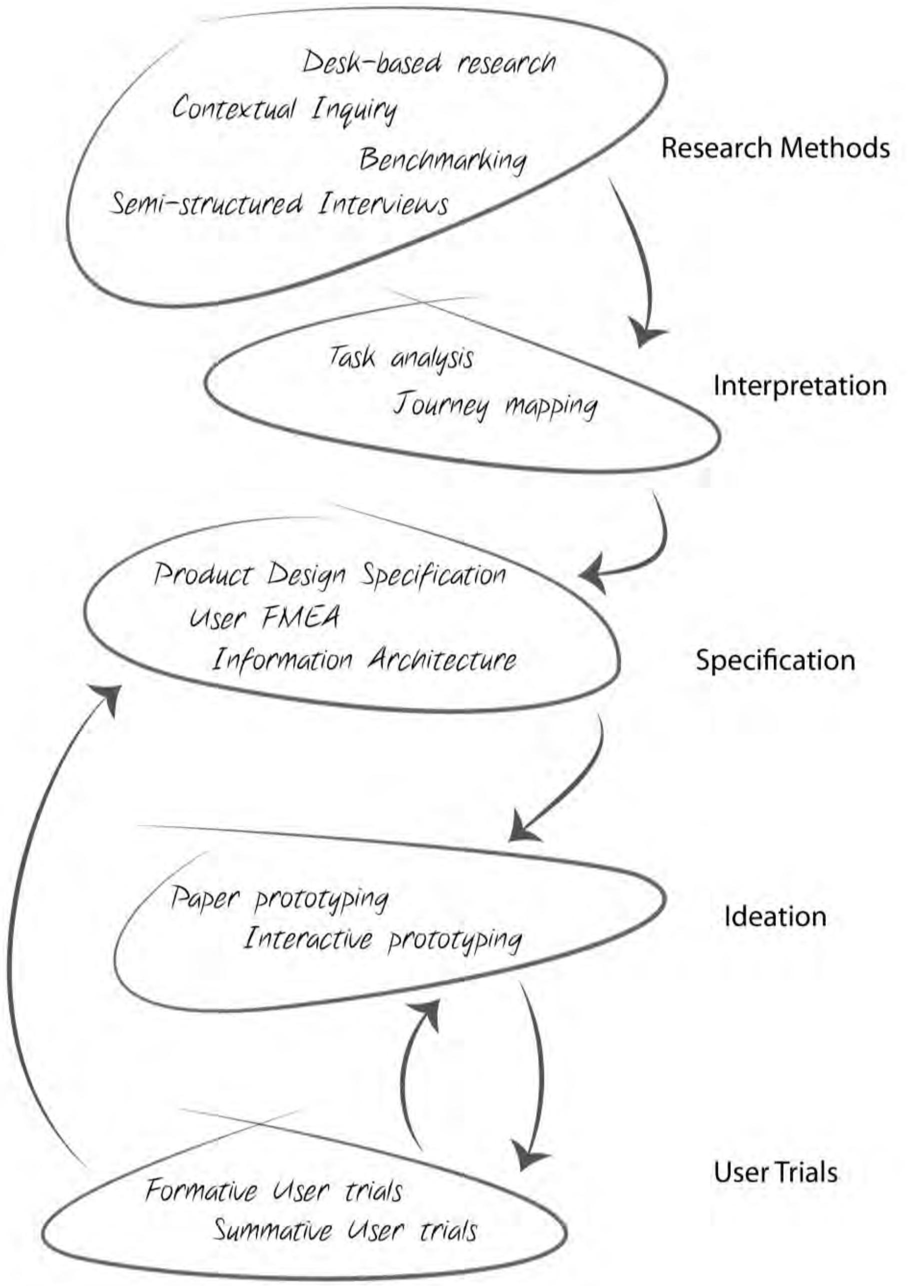


Figure 8: An overview of the typical UCD processes used at PDR

This initial data is gathered into a report in which the different data types are analysed, creating insights into the product area from the perspective of the potential user. Some future-thinking design projects which investigate an area of design could even finish with this report, from which a design brief can be established.

Next, interpretation techniques can be used to capture the users' perspective in a way that can be discussed by the design team. Journey mapping originated in service design and describes the journey of a user by representing the different 'touchpoints' that characterize an interaction with the service or product (Temkin, 2010). Personas are fictional characters created to represent various user types within a targeted demographic, attitude and behaviours that might use a product (Goodwin, 2009, p. 229). Personas are useful in considering the goals, desires, and limitations of users in order to guide decisions about a service, product or interaction space such as features, interactions, and visual design.

Specification methods are the more formal techniques that specifically capture user orientated data, this is critically important in medical products which require proof that a user-centric design process has been applied to reduce potential user errors once the product is released.

Ideation techniques realise product and interaction concepts in a way that can be discussed with potential users. A concept can be realised in number of ways depending on how developed the concept is. Prototypes are the primary form of ideation and these will be covered in detail in the next section. Another technique is storyboarding (Rosson & Carroll, 2002), which describes how the user and product might interact in a given scenario; this can be a powerful tool creating discussion around both the products' functionality, its interaction and the scenario proposed.

Once the concept has been realised, users can be consulted in a variety of ways depending on the stage in the design process. User testing is a way of gathering empirically based information and applying it to the design of things both in the development and evaluation of designs (Sanders & McCormick, 1993, p. 23). The term user 'testing' has come under criticism because it is not the user that is being 'tested' but the product, therefore the term user 'trial' will be used in this thesis. Techniques employed in user trials will be covered in more detail in the next section, but include usability, task analysis, semi-structured interviews and in-context studies. Additionally, many of the techniques from the research phase can be used; for instance a video diary could be undertaken if a working prototype was created.

Formative and summative studies are both forms of user trial based on a product or a prototype of a product. Formative user studies tend to be smaller studies which are intended to inform the design process, and can take place throughout the iterative design process (Tullis & Albert, 2008). Summative user studies tend to involve a larger number of participants and are intended to evaluate how well a product meets its objectives (Tullis & Albert, 2008). Summative user trials might collect more quantitative data such as task times, completion rates and satisfaction scores.

The diagram depicts that the process described here is iterative so that insights from users are constantly being fed into the design process. Ideation and user trials are the most common phase to build, evaluate and learn but these phases also inform the product design specification. Additionally, the research methods could be utilised within the user trial phase observing prototypes in use.

Designers need meaningful feedback from users and the prototype provides the basis of communication, because sketches can be difficult to understand by non-professionals (Stacey, Eckert, & McFadzean, 1999). Therefore the role of the prototype in supporting the UCD process as a physical manifestation of that process is of specific interest here.

2.3.7 Design Considerations Specific to the Computer Embedded Product

This section began with a brief history of HCI and Industrial Design. In bringing HCI specialists and Industrial Designers together, the Industrial Designer needs to understand more about programming and electronics, which fall outside their traditional skill set, in the recognition that they are no longer designing solely what the product looks like but also “how it behaves” (Moggridge, 2007, p. xvi). In addition, the HCI specialist needs to recognise the difference between programming for a desktop computer and a computer embedded product (Bergman & Haitani, 2000, p. 2).

Goodwin (2009) proposes five roles that the design team of computer embedded products should include, these are: two types of interaction designer (the ‘generator’ and the synthesiser’), a visual designer, an Industrial Designer and a team leader. The two types of interaction designer arise in recognition that computer embedded products change state over time; therefore the interaction designer needs to understand the 'narrative' of the design, in addition to the structure of the interface. Goodwin advocates two interaction design roles, that of the Generator and the Synthesiser, where the generator leads the visualization of system behaviour (its information architecture), and the synthesiser analyses concepts from a narrative point of view. Information architecture is concerned with the structure of content; how best to organise and label content so that users can find the information they need (Saffer, 2010).

Another notion useful to designers of computer embedded products is the mental model. The notion of mental models originated in cognitive psychology as internal constructions of some aspect of the external world that are manipulated, enabling predications and inferences to be made (Craik, 1967). A mental model is the knowledge that people develop of how to interact with a system. The user does not necessarily need to know how a product works, just have enough understanding to be able to use it. This does not require deep knowledge of the technology behind the product, just a 'story' of how it works. However, many people do not understand how computer embedded products work (Rogers, Sharp, & Preece, 2011), therefore their mental models are often incomplete, easily confusable, or based on inappropriate analogies and superstition (Norman, 1983). For example, if you are in a rush you might set an electric oven to higher than the desired temperature thinking it would heat up 'faster'. Many people base this on a mental model that suggests 'more is faster' but in the case of an oven, the heating element will heat at the same rate regardless of the temperature it is set to, the temperature setting instead relates to the final temperature it will achieve. Designers of computer embedded products need to consider the information architecture of the product to aid the user in creating an accurate mental model of how to interact with the product.

The requirement to change state in response to a user necessitates the involvement of a programmer during design because many designers lack the competency to effectively implement their ideas in software (Rosson, Ballin, & Rode, 2005). Designers might present good quality static images to a programmer but without any time-based functionality of the subtle flow characteristics, concepts might get 'lost in translation' unless the designer and programmer work very closely.

Rapid iterative development, as proposed in Section 2.3.3, requires evaluation of ideas through prototypes in order to ensure sound decision making before a product moves towards production when problems found can be extremely costly to rectify. Yet for computer embedded products meaningful prototypes are difficult, the Industrial Designer can mock-up the physical form yet this is almost meaningless without its digital intentions. New interactive design methods, techniques and tools are needed to help the designer externalise and communicate their ideas (Branham, 2000). Many research groups have tried to address these needs. The techniques and tools developed will be discussed in Section 2.3.8.

2.3.8 External Representations of the Product Design Process

Römer *et al*. (2001) divided the external representations of the design process into four groups, that of sketching, CAD modelling, and both simple and complex prototypes.

2.3.8.1 ***Sketches***

Sketches are a style of drawing; they are 'thinking drawings'. They are two-dimensional and inherently fast and ambiguous (Gaver, Beaver, & Benford, 2003). One of the key purposes of sketching in concept design is to provide a catalyst to simulate new and different interpretations (Buxton, 2007):

> *"… designers do not draw sketches to externally represent ideas that are already consolidated in their minds. Rather, they draw sketches to try out ideas, usually vague and uncertain ones. By examining the externalizations, designers can spot problems they may not have anticipated. More than that, they can see new features and relations among elements that they have drawn, ones not intended in the original sketch. These unintended discoveries promote new ideas and refine current ones." (Suwa & Tversky, 2002)*

2.3.8.2 ***CAD models***

The next group from the classification of external representations by Römer *et al*. (2001) is Computer Aided Design (CAD) models. These are models created on a computer and will only exist in a virtual world, these models will be discussed in greater detail in the next section. CAD models have the same limitations as sketches in relation to their physical interactions: *"in general, people are good at experiencing 3D and experimenting* *with spatial relationships between real-world objects, but possess little innate comprehension of 3D space in the abstract. People do not innately understand three-dimensional reality, but rather experience it"* (Scali, Shillito, & Wright, 2002).

2.3.8.3 ***Prototypes –Simple & Complex Models***

Both simple and complex models from the definition of external representations of the design by Römer *et al*. (2001) are forms of prototypes. As defined by the Merriam-Webster Dictionary, a prototype is '*an original or first model of something from which other forms are copied or developed*' (Merriam-Webster, n.d. (b)). A 'prototype' literally means 'first of a type'. When viewed in terms of mass production, this notion makes sense for a model that is produced in advance, exhibiting all the essential features of the final product and used as a test specimen and guide for further production (Floyd, 1984). In the International Standard for Human-centred design process for interactive systems (BS EN ISO 13407:1999) a prototype is defined as a representation of all or part of a product or system that, although limited in some way, can be used for evaluation.

In the classification by Römer *et al*., simple models equate to low-fidelity prototypes and complex models equate to high-fidelity prototypes. Low-fidelity prototypes are covered in more detail in Section 2.4.3 on page 53.

The use of the term prototype can, and is, interpreted in many ways, and the exact use of the term 'prototype' differs across the product design and HCI disciplines. In product design, Pugh (1991, p. 175) discusses prototypes being used both at the beginning of the design process and at the later stages, but emphasises that these prototypes have very different functions. The early prototypes are used to gain a clear conceptualization of the basic issues of the design, whereas the later stage prototypes are to establish the technical feasibility of the product.

Ulrich & Eppinger (2012, p. 294) define prototypes by the four purposes they regard them as being used for: learning, communication, integration and milestones. In addition, they describe prototypes as having two dimensions: physical to analytical, and comprehensive to focused. Physical prototypes are tangible artefacts created to approximate the product whereas analytical prototypes represent the product in a non-tangible, usually mathematical, manner. Comprehensive models implement all of the attributes of the product whereas focused models implement just a few attributes. Lim *et al*. (2006)

describe two characteristics of a prototype: as an incomplete portrayal of a design idea and a manifestation of a design idea or ideas.

Lennings *et al*. (2000) describe five categories of prototypes as being based on their functions:

- **Shape** models –these models represent the outer appearance of the design
- **Functional** models –to test the functioning of some part of the design
- **Physical behaviour testing** models - used to simulate certain behaviour of (a part of) the design, like strength or stiffness
- **Presentation** models –to present the design to an outsider, often these will be highly finished
- Models for **stimulating group discussion**

This thesis focuses on the initial stages of the design process, and Lim *et al*. (2008) regard the prototypes used in these stages as being those to enable design thinking, they are the means by which '*designers organically and evolutionarily learn, discover, generate, and refine designs'.* This can be any of the prototype functions defined by Lennings *et al*., but more likely only the learning and communication prototypes defined by Ulrich & Eppinger.

Buxton argues that the prototypes used in these early stages should be classed as sketches "*essentially the investment in a prototype is larger than in a sketch, hence there are fewer of them, they are less disposable, and they take longer to build"* (Buxton, 2007, p. 139). Figure 9 shows Buxton's view of the difference between a prototype and a sketch, yet Schrage argues that:

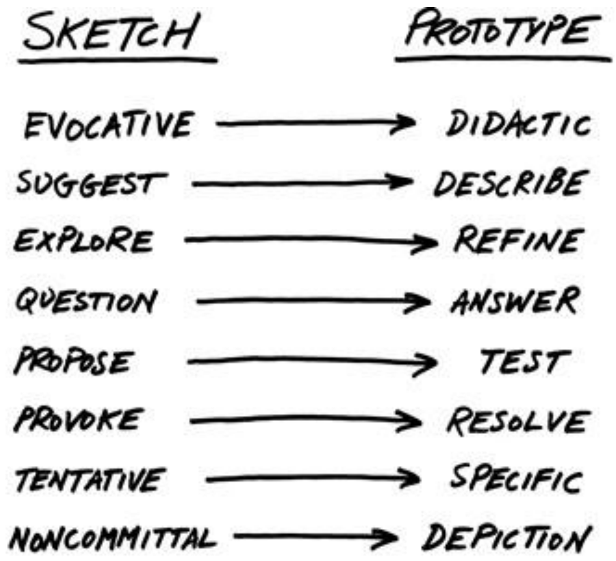


Figure 9: Buxton's view of prototyping and sketching

> *"Prototypes and simulations can do more than answer questions; they can also raise questions that have never been asked before. Playing with a prototype can stimulate innovative questions as surely as it can suggest innovative answers. The best and most powerful models are provocative, and the unexpected questions that a model raises are sometimes far more important than the explicit questions it was designed to answer." (Schrage, 1999, p. 76).*

Baxter agrees with Buxton's definition, and states: *"the general rule in developing prototypes is 'only do so when necessary'."* (Baxter, 1995, p. 285). This view is reflected by some of the representations of the design process as discussed in the previous section, in which prototyping is seen as part of the detail design and 'fit for purpose' testing.

The difference seems to stem from terminology, the prototypes that Buxton and Baxter refer to are those created during the later stages of the design process. These take the meaning of prototype literally as the 'first of the type' of the final design, created to establish the technical feasibility of the entire product (Pugh, 1991).

The research presented in this thesis refers to those prototypes which are described by Lim *et al*. as 'thinking enablers'. Many authors argue for the increased use of prototypes in human centred design especially in the initial stages of the design process (Buxton, 2007), (Kelley, 2001), (Schrage, 1999). Because a lower investment is required, both in time and money, low-fidelity prototyping is a really fast way to make good quality decisions about the potential of an idea:

"At this stage these early prototypes and studies are not tests of the design; they are tests of your understanding of the issues." ... "When the development team and the customers are happy, then it is time to translate the mock-ups into design specifications." (Norman D. , 1999, p. 187)

Physical prototyping has quite an obvious potential in respect to physicality, therefore questions around the appropriateness of prototypes will be explored in more detail in the remaining part of this literature review.

2.3.8.4 ***Tools and materials used in the Product Design Process***

The design process is a theoretical model adapted to suit the individual needs of a project, design team or company philosophy. Physicality is part of the tools and techniques used within this model, from sketching ideas through making prototypes to detail design and beyond. Therefore, an understanding of how tools are, and have been, used during the design process is important in understanding how physicality affects product design.

For this research, the tools and materials used in the product design process (PDP) can be broken down into two types: those of the physical world, for example sketchbooks and blue foam, and those of the digital world, involving computers.

In respect to Römer *et al*.'s external representations of the design process, physical tools will be used for sketching and for the creation of both simple and complex models. Tools and materials used for sketching tend to be paper based, plus a pencil/pen/marker pen/crayon. Simple and complex models can use anything from card, blue foam, wood and modelling clay. Extending this, designers use any number of components salvaged from another product or another purpose. At the leading design consultancy IDEO, Kelley (2001) describes an area they put aside for the collection of potential materials. In fact *"Industrial Designers tend to create models themselves and use whatever materials and tools are available and convenient."* (Gribnau, 1999). Trudeau (1995) describes many fabrication techniques for physical models in his book, Professional Model Making.

Modelling clay has a long history of use in the 3D design process by many designers 'according to the type of product and company practice' (Bordegoni & Cugini, 2005), most notably by the automotive industry (Verlinden, Wiegers, Vogelaar, Horváth, & Vergeest, 2001). But even more prevalent in design today is blue foam or Styrofoam. Dyson, known for its vacuum cleaners, makes sketch models using cardboard and foam, and Rodd Industrial Design use foam in the assessment of form and ergonomics (Evans, 2005).

But more of the tools used in design are becoming digital. Computer aided design (CAD) became readily available as a solid modelling tool in the late 1980's (Dieter, 2000, p. 269), and is now utilised across most of the product design industry. Römer *et al.* (2001) revealed in their survey of practising designers that 93% of the 106 participants required a computer based output. For industrial processes, the "*latter stages of product development such as … manufacturing planning requires the product model to be available in electronic form*" (Scali, Shillito, & Wright, 2002).

CAD programs can be used at specific times in the design process, and, as technological sophistication increases, so too do the capabilities of CAD programs. CAD programs can be used to sketch out initial design ideas in a rather loose way (for example, Rhinoceros 3D), or by creating and manipulating the surfaces of designs (for example, ICEMSurf), or by building a solid model of an idea (for example, Solidworks, CATIA, ProEngineer). Computers can also aid the design process in the construction of two-dimensional representations of a design, either by the use of graphics programs such as Adobe Photoshop and Illustrator or through the direct 'rendering' of images from CAD packages. In addition, computers can be used to aid in manufacturing decisions by testing out intended designs in a virtual world to see how they would react to physical manipulations (Finite Element Analysis), or how plastic might flow if that part was to be injection moulded.

These are just a few of the ways computers are being used in product design at the moment. All of these programmes are run on desktop or laptop computers where the traditional input products are a mouse and keyboard. Therefore the human element of the design can be lost, for example, a designer can zoom into a small detail of the design and neglect the overall design. Yet the computer has undoubtedly transformed the way design is conducted for the better (Narayan, Rao, & Sarcar, 2008).

A number of products that help the designer interact with the computer in a more 'natural' way have been developed; the 'space mouse' can be used to rotate and control a 3D virtual model in most CAD packages. Graphics tablets can be used to help the designer 'sketch' on a computer in a manner more akin to traditional pencil on paper sketches. The development of better touch screens has begun to supersede graphic tablets by enabling the designer to sketch directly on the screen. And even more recently, developments in haptic technology have enabled the designer to directly manipulate virtual clay using, for example, Sensable's Phantom haptic arm (now Geomagic Touch Haptic Device) with the virtual modelling software 'Freeform' by Geomagic as shown in Figure 10 below.

'Freeform' is based on the principle of a virtual piece of clay. This can be manipulated through the Geomagic Touch haptic device with which the user can 'feel' the surface of the 'clay'. Freeform can be operated through the Geomagic Touch Haptic Device directly on the model or through the use of its more traditional CAD functions such as the creation of sketch lines.

Figure 10: Sensable's Phantom haptic arm (now Geomagic Touch) with 'Freeform' displayed on screen

But are these developments enough? Artists and designers find that input products frequently lack the haptic and force feedback they would expect from conventional crafting tools (Shillito, 2004) which can lead to dissatisfaction and frustration. Practitioners often rely heavily on digital tools that lack fine sensitivity to pressure and gesture and remark that the complex neuromuscular potential of fingers and thumbs is rarely exploited in current technology (Treadaway, 2007).

Finally, computers have affected the way in which prototypes can be created. Additive manufacturing (AM), which is also known as Rapid Prototyping (RP), is a technique by which one-off prototypes can be constructed with the precision generally associated with mass-production techniques. AM models are first created in a CAD package and AM techniques will breakdown the CAD model into layers and then build those layers, one at a time, in the physical world. The Fused Deposition Modelling (FDM) machine will extrude the layers in plastic. The Stereolithography (SLA) technique uses a laser to cure resin layer by layer. And the Selective Laser Melting (SLM) machine uses a laser to fuse metal powder together, one layer at a time. AM models are often created at the later stages of the design process due to the more precise and less exploratory nature of the models plus the expense of the machinery (Broek, Horváth, & Lennings, 2000).

2.3.9 Summary

This theme provides an overview of the complex skills involved in the design of computer embedded products. It highlights the different histories and skills needed, specifically the skills of the Industrial Designer (physical aspects of the product) and HCI specialist (digital aspects of the product).

The change from technology-driven to task-driven products has caused the design process to become more user-centred. There are a variety tools and techniques of the user centred design process, these can be grouped into; generative research methods, interpretation methods, specification methods, ideation techniques and user trials.

Physicality is identified as being fundamental to the tools and techniques used during the design of computer embedded products, from sketching ideas through making mock ups and prototyping. The Contextual Study develops the understanding of physicality in relation to computer embedded products by focusing on the design process itself.

2.3.10 Contextual Study - the 'Torch Project'

The 'Torch Project' was a two-phase project which was designed to explore the role of physical tools and materials during the concept stage of the product design process.

The first phase can be found in Appendix 2, this is a detailed account of a reflective investigation of the design of a handheld torch.

Phase two was a conducted as a design exercise and the following publications can be found in Appendices 3 & 4.

> Ramduny-Ellis, Hare, Dix, & Gill (2008), **Exploring Physicality in the Design Process.** In the proceedings of the Design Research Society Conference 2008.

Ramduny-Ellis, Hare, Dix, Evans, & Gill (2009), **Physicality in Design: an exploration.** In The Design Journal 13(1), pp 172-189.

The literature review revealed that when computers began to be used within the design process they bought virtuality, a stage that cannot be physically 'felt' or explored. Technologies which could help to address this highlighted by the literature review included 3D scanning technology, Additive Manufacturing (AM) processes and haptic control of Computer Aided Design programs. The first phase of the 'Torch Project' investigated the use of these technologies through a reflective investigation in the design of a handheld torch and compared them to typical Computer Aided Design (CAD) approach. The differences in the prototypes created seemed to be as a result of the clay versus sketch approach. This insight drove the second phase of the 'Torch Project' which explored the impact of different materials during the concept generation stage of the design process.

2.3.10.1 ***Phase 1 – an overview of the study***

A design brief was chosen based on one of the most basic of interactive products, a torch. The project focussed on the conceptual stage of the design process where the difference of material and tool usage has the greatest potential impact.

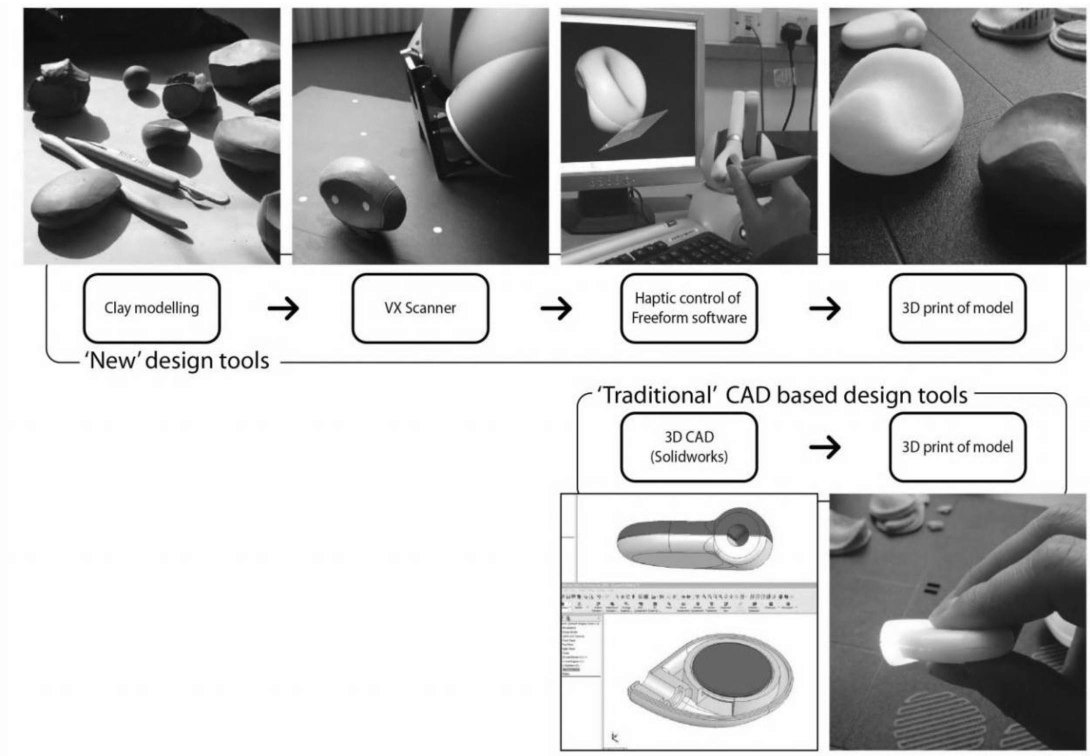


Figure 11: An overview of the two approaches to the brief for the reflective investigation –using 'new' and 'traditional' design tools

To enable a comparative reflection, this brief was approached in two ways; the 'new design tools' approach used 3D modelling tools identified in the literature review as enabling a more 'hands-on' approach; and the 'traditional' approach was CAD-based (Figure 11). The fundamental difference between these two approaches is the way in which the concept is explored; the 3D modelling tools used in the hands-on approach enables exploration of ideas physically with potters' clay whereas concepts were explored with pencil on paper sketches in the traditional approach.

2.3.10.2 ***Conclusion of phase 1***

The inclusion of the 3D scanner in the 'new' design tools approach enabled the use of clay modelling which provided a 'hands-on' approach to constructing the physical models. When creating models with clay, the designer can quickly explore the relationship of the model to the hand and body. Resultant models of the 'hands-on' approach were more organic than those of the 'traditional' approach; these tended to be based around the components specified for the product, with less consideration to the way in which it would be held.

However, the use of the scanner had its own limitations. The fundamental limitation in this study was the inaccuracy when scanning palm-sized objects or smaller. Perhaps alternative scanning technology could be used such as the Photon 3D Scanner, for which, the small object is placed on a rotating platform to be scanned.

In addition to the limitations of the scanner, the Freeform modelling software was not found to be as natural as expected. There seemed to be two reasons for this. Firstly, it was difficult to gauge the position of objects on screen which meant that continual rotation of the model was necessary to avoid unintended distortion through the depth plane. Secondly, it felt very unnatural manipulating clay through a product that is held in the same way as a pen. Both of these limitations have the potential to be overcome with practice. Animators, for example, use this combination to create detailed models suggesting that it can be learnt given enough practice. However, given that the intention of such tools is the use of haptics to provide 'true-to-life sensations' (Geomagic, n.d.), the need for considerable practice seems counter-productive. When comparing the interaction with Freeform to the use of potters' clay, a potential cause can be seen. For the potters' clay, both hands were used to model the clay, this included the palm, fingers and finger tips, in addition to specialist clay modelling tools. When using Freeform, this interaction is conducted through a pen-like product with force-feedback. Despite the ability to choose tools and tips shapes in Freeform, the feel of interacting through a pen-like product did not replicate the whole hand interaction of modelling clay in the physical world.

Evans (2005) conducted a case study using the Phantom haptic product with the Freeform software. He used the Phantom for the creation and manipulation of the design and found that the haptic part of the software/hardware combination proved 'ineffective'. Evans proposed that this was because the tactile interaction resulted in "*a degree of undulation on curved surfaces that was considered unacceptable for both rendering and downstream tooling operations*". Although the quality of these surfaces could be improved through the use of traditional CAD techniques, such as guide lines, this was moving away from the use of tactile feedback and therefore considered inappropriate in both the study by Evans and the study presented here.

Two points for 'hands-on' interaction are identified in this investigation; the use of clay to explore ideas and the Geomagic Touch haptic device to manipulate and refine ideas. As such, a design exercise was created for Phase 2 of this study to explore the use of modelling clay and other physical materials. The use of Geomagic Touch haptic device in providing the second opportunity for hands on interaction would seem ineffective based on this study and that of Evans described above. Yet it is inconclusive as to whether this is due to technological limitations or the interaction style. Further research would need to be conducted in this area to explore what causes this interaction to feel unnatural.

2.3.10.3 *Phase 2 –an overview of the study*

A design exercise was constructed to explore the influence of a variety of materials in the initial stages of the design activity by deliberately restricting the use of design tools and materials during a short design brief. Analysis sought to discover how the physical properties of the allocated design tools and materials impact issues such as the number and novelty of design ideas and the kinds of designs produced.

Materials were chosen to be accessible for a wide range of people which would not require any specialised skills. The choice of materials also reflects traditional design practice and covers a range of properties such as, two-dimensional vs. three-dimensional, manual vs. cerebral, malleable vs. constrained. The 'kit' of materials was:

- paper and pencils
- card and glue
- modelling clay (commonly known as plasticine).

2.3.10.4 ***Conclusion of Phase 2***

The study concluded that it was not possible to identify a direct relationship between the type of material used and the approach each team adopted. The results do not support simplistic conclusions such as 'physicality promotes creativity', or even the opposite.

The choice of material clearly has an impact, but the individual skills and background of the participants are equally important. Materials can constrain people, but they can also inspire creative design. But equally, less tangible expression in discussion and written form seems to allow breadth of exploration.

2.3.10.5 ***Conclusion of the 'Torch Project'***

Phase 1 and 2 of the Torch Project suggested that the use of materials has an effect on the type of concepts created. However, these effects are not simple to define but rather complex and dependent on the design team involved. Suggestions to take this work forward included an observational analysis of the way practising designers use prototypes and physical models in a commercial context and further investigations into the use of the newer technologies of 3D scanning and haptic interface products.

Note that as of January 2014 a further study has begun. In this study, four established artists have been invited for a sequential 8 week residency at PDR (due to end in July 2014). During this time they will be instructed on the use of the haptic interface product, the Additive Manufacturing technologies and 3D scanner as appropriate to each artist. Each artist will then construct a piece of artwork with these technologies. The artists are required to complete a reflective journal of their journey. There are several research aims of this project although the theme being pursued as a result of this PhD research is an investigation of the effect these new technologies have on the way the artists create work.

2.3.11 Conclusion

This theme covered the disciplines and techniques involved in the creation of computer embedded products. These products are physical devices and therefore physicality is an intrinsic part of the design process. However, computers are increasingly used in the design process which brings 'virtuality', a stage that cannot be physically 'felt' or explored.

A number of tools were identified to have the potential to bring physicality 'back' into the design process. These tools included 3D scanning technology, Additive Manufacturing (AM) processes and haptic control of Computer Aided Design programs. Phase 1 of the contextual study, the 'Torch Project' (Appendix 2), suggested that the tools identified by the literature review were not yet at a 'natural' level. Given that the intention of the use of haptics is to make 'true-to-life sensations', the need for considerable practise with the tool seems counter-productive. Yet there was a clear difference in the way in which the design materials influenced the final design. The standard CAD prototype had a 'functional' feel and was highly symmetrical. In contrast, the prototype from scanned clay was more 'organic'.

The investigation questioned the impact different design tools have on the outputs of the design process, which was explored by the design exercise of the contextual study, presented in Phase 2 (Appendices 3 and 4). This study revealed that it was not possible to identify a simple relationship between the type of material used (and the relative physicality of that material) and the prototypes created, however the choice of material did have an impact on those prototypes.

The review of literature and contextual study identified that physicality is intrinsically part of the way computer embedded products are designed. Most notably for this research, it will have an effect on the way designers realise ideas through models and prototypes, even if that effect is not simple to define. These models and prototypes are used throughout an iterative design process and therefore the research question will address the way in which designs are realised.

2.4 Theme 3: Physical Manifestations of the Design Process

To incorporate interaction in a prototype the model needs to be responsive to a user's action, or at least give an appearance of being responsive. There are a number of research groups both from industry and academia who have produced toolkits and methods for creating interactive prototypes. The toolkits and methods range from two-dimensional screen-based techniques and 'faking' the feedback of the product, to techniques making models truly interactive through the addition of electronic modules in the prototypes to “deal with time, phrasing and feel” (Buxton, 2007, p. 139).

So called ‘quick and dirty’ prototypes are a fundamental tool for many for these techniques, and interactive prototypes can be used to explore the digital considerations within the physical form of computer embedded products.

Finally, the notion of fidelity is discussed in relation to interactive prototypes of computer embedded products.

2.4.1 Interactive Prototyping without Electronics

There are some notable interactive prototyping techniques that do not require electronic integration. These either forgo the physical model entirely, or completely ‘fake’ the feedback.

2.4.1.1 *Screen Based Prototypes*

Screen based prototypes are graphic representations of both the product and its graphical interface shown on a computer screen as shown in Figure 12. The user interacts with the prototype by touching the screen to operate the interface. For the prototype shown in Figure 12 the user presses the 5 button cluster to move the active screen element. The benefit of this approach is that it can be created relatively quickly, yet there is no tangible model to hold and there are no physical buttons to interact with.

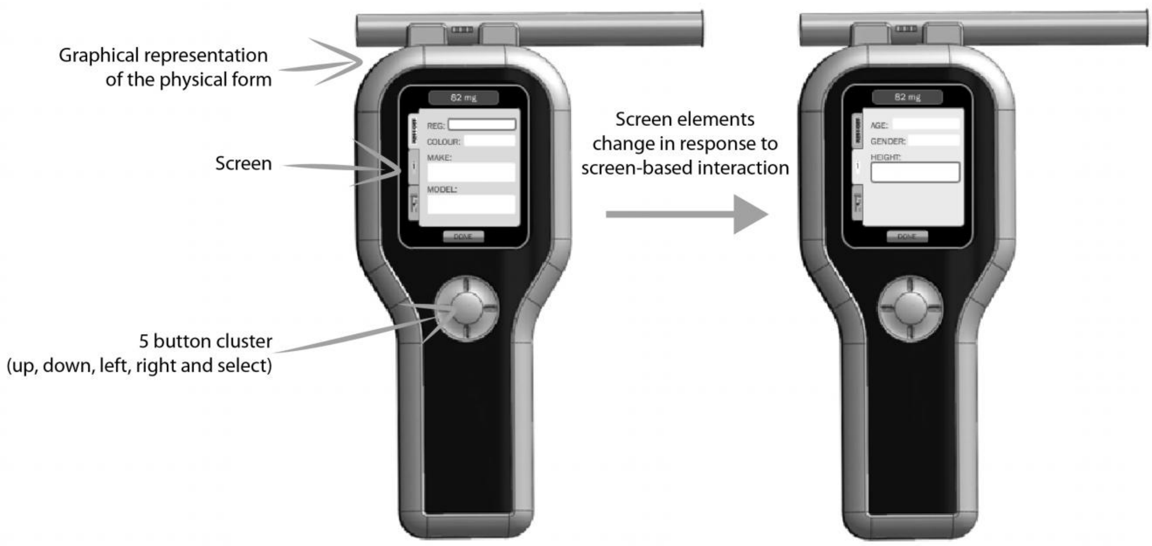


Figure 12: Example of an on-screen interactive prototype

This is a common technique used in industry (Pering, 2002). Culverhouse (2011) undertook a detailed investigation of 6 companies and found that 3 regularly used on-screen prototypes. Although this is a relatively small sample it demonstrates that on-screen prototypes are still used 9 years on from Pering's original study. Yet there has been little research undertaken to prove that these prototypes give accurate feedback on the complete product with its physical and digital forms.

2.4.1.2 *Faking it*

Faking interactions can produce extremely powerful experiences. The prototype is completely physical and will rarely have functional embedded electronics. These methods can be extremely quick and low cost but will often require facilitators to stand in as the computer; therefore they are not 'ready to go' prototypes.

Paper prototyping (Snyder C. , 2003) uses sticky notes or screen sketches to represent the on screen functions of the product. The user will hold a physical mock-up of the form (for example a foam model) and the facilitator will adjust the screens depending on how the user interacts with the product. The very nature of a paper prototype invites the user to give feedback on fundamental issues of the design (Holtzblatt, Burns Wendell, & Wood, 2005, p. 248). Three-dimensional paper prototypes are a combination of user interface design methods and Industrial Design methods. They represent the basic functionality, but also the basic design ideas in three dimensions (Sade, Nieminen, & Riihiaho, 1998).

Wizard of Oz prototyping (Buxton, 2007) uses another person to simulate the digital functions of a product. The classic example of this is IBM's simulation of voice recognition for typing documents (Gould, Conti, & Hovanyecz, 1982). For this prototype the user would speak to the product and a person, hidden elsewhere, would type what the user

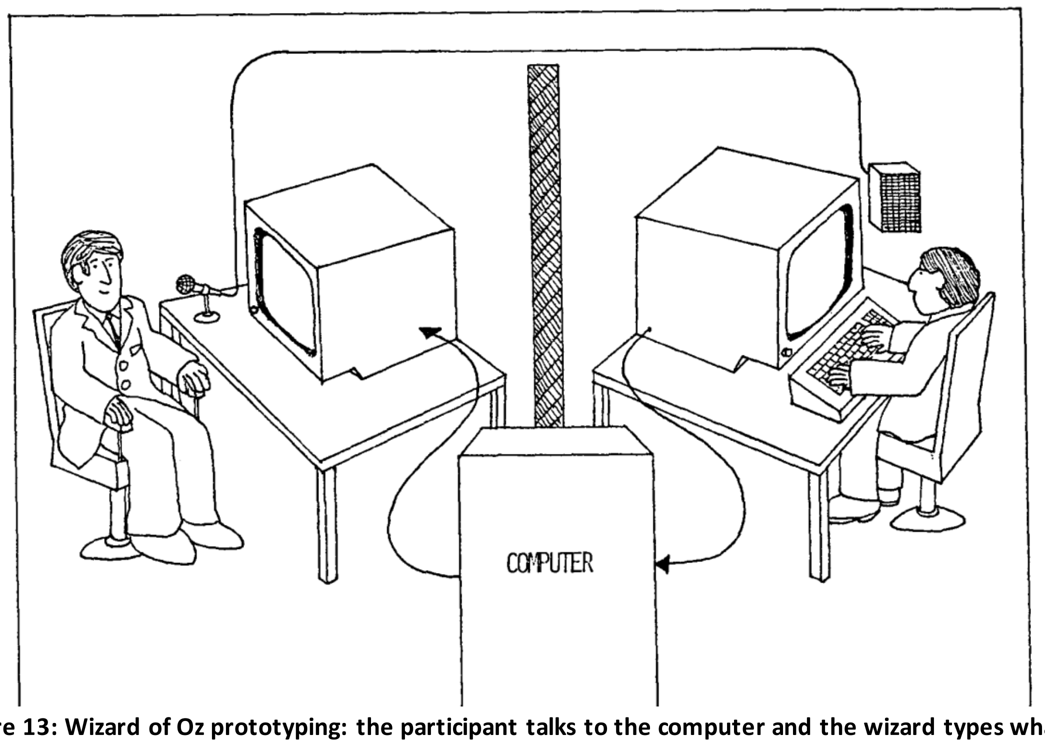


Figure 13: Wizard of Oz prototyping: the participant talks to the computer and the wizard types what is heard, this appears on the participants monitor and appears as voice recognition

has said, making it seem like the product is responding as shown in Figure 13.

2.4.2 'Smart' Interactive Prototyping

'Smart' prototypes will react directly to the users' interactions. They require some form of electronics to be integrated into their physical form which will directly control the digital functions of the product. There are several toolkits available, these include Ardunio (Burleson, Jensen, Raaschou, & Frohold, 2007), the IE System (Gill, 2013), Phidgets (Greenberg & Fitchett, 2001), DTools (Hartmann, Klemmer, Bernstein, & Mehta, 2005), Pin and Play (Villar, Lindsay, & Gellersen, 2005), Littlebits (Bdeir, 2009), and Makey Makey (Silver & Rosenbaum, 2012) to name a few.

Most toolkits work as translators, feeding interactions with the product into a PC in a usable form. Phidgets, for example, are 'physical widgets' which can be integrated into the physical product. *"Phidgets are a set of 'plug and play' building blocks for low cost USB sensing and control from your PC".* They are complete electronics solutions for various interactions that can be wired to a PC running the 'Phidgets library' which will translate the users' interactions into a form understandable by the computer. The Phidgets library is a set of mini-programs dedicated to the Phidget and the software used. The benefits of this system are that complex interactions can be constructed fast, but this means the electronic components have defined physical dimensions and therefore they may not fit the physical constraints of the design. Also, despite the Phidgets library being freely accessible, the level of required programming skills is often beyond that of a product designer.

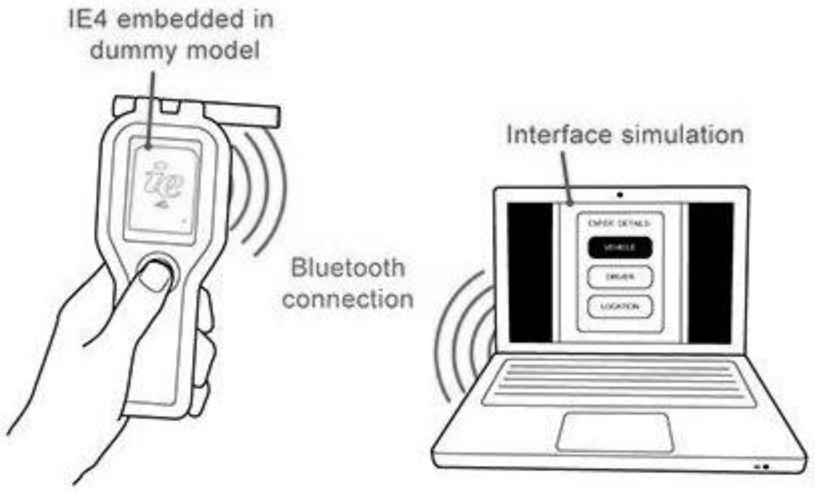


Figure 14: The IE System

The IE System (Gill, 2013) works by connecting the physical model embedded with standard switches to a computer running the Graphical User Interface (GUI) via the *IE Unit* (Figure 14). The IE Unit translates standard button presses on the physical model into keyboard inputs. Any software program that can be triggered by keyboard presses can be used to model the GUI. When a user activates a switch in the model, the computer responds to a perceived keyboard input and a keyboard triggered GUI is activated. The IE system can use any switch and any material to make a model, this means that ‘to-scale’ representations of a design can be constructed. It uses ‘dumb’ models, electronics and any software making the system accessible to many people. The limitations of this are that it uses essentially digital signals. Analogue signals such as sliders and accelerometers are not easy to simulate.

Smart prototyping techniques can be used and adapted depending on which interactions need to be embodied by the prototype. Some are limited by the physical size of the toolkit, others by their accessibility, or by the skills needed to control the interactions.

Interactive prototypes are based on five fundamental elements as shown in Figure 15. A **physical model** needs to be created upon which **interactions** can be added, then a **prototyping platform** is required to translate user actions into an input that can be used by the **prototyping software** which will then visually show the interface through the **hardware**.

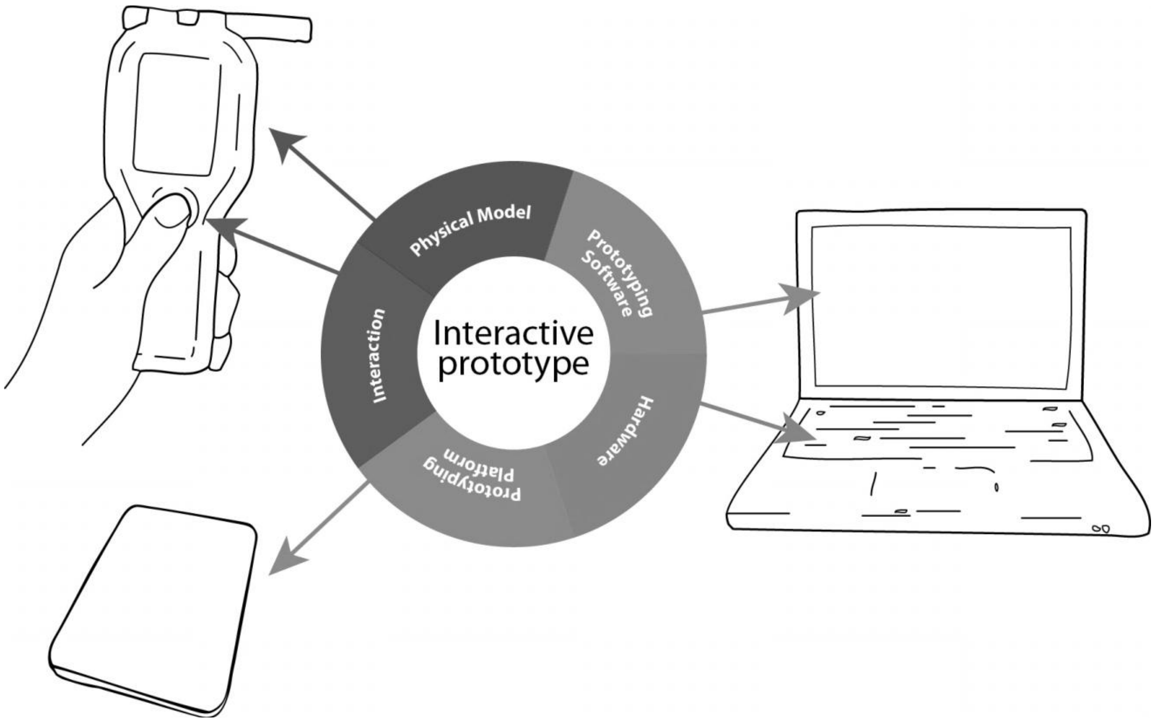


Figure 15: Five elements of an interactive prototype

There are a number of **prototype software tools** available to construct the interface of prototypes such as ‘click throughs', wireframes, animations and coded interfaces. One online source lists over 100 different software tools (Goltz, 2012). Another list from user experience specialists’ ‘Adaptive Path’ lists 42 software prototyping tools (Harrelson, 2009). Tools include Axure, Adobe Flash, Microsoft PowerPoint (Kelly M. , 2007) and Nokia’s Flowella (Nokia, 2011).

If the product being designed is not mouse and keyboard driven or touchscreen, a further set of tools, called **electronic prototyping platforms**, are used to translate user actions into an input that can be used by the prototyping software tool. The Arduino (Burleson, Jensen, Raaschou, & Frohold, 2007) is probably the most prolific of these tools, but many others exist including the IE system (Gill, 2013) and Phidgets (Greenberg & Fitchett, 2001) which were covered earlier. EventHurdle (Kim & Nam, 2013) translates hand-held, remote and touch gestures for use with Flash ActionScript. With the exception of the IE System, each of these platforms is based on a dedicated computer program within which code can be written to translate inputs into outputs.

The electronic prototyping platform and prototype software requires some form of computer (**hardware**) on which to operate, this could be a laptop or desktop computer, a tablet or 'palmtop', or even a microcomputer such as the Raspberry Pi.

Unless the product is purely touch screen or has a gestural interface, a prototype will also need **electronic interface products** such as buttons, dials and sliders. For some systems, these can be off the shelf (IE System) and for others dedicated inputs are required (Phidgets).

Finally, the prototype needs a **physical form**. A number of techniques exist for creating a model with which the user can interact from very rough materials such as paper and cardboard, through model board and blue foam to rapid prototype techniques such as FDM and SLA which require prototyping machinery.

The Raspberry Pi (RPi) is a microcomputer which runs a modified Debian (Debian, n.d.) operating system such as Raspbian. Raspberry Pi aims to make programming accessible to everyone, especially children "all over the world". There has been a huge response to the launch of this microcomputer, with a plethora of information and open source code available through forums. Physical accessories can be bought specific to the needs of a project, for example a camera or light sensor. As such, the Raspberry Pi provides the prototyping platform, software and input product within one system which could potentially be entirely embedded in a prototype.

The combination of software, electronic prototyping platform, interactions, hardware and physical form used to create a prototype will depend on the demands of the prototype being created, and the skill of the person constructing the prototype. Many platforms are heavily dependent on the various coding languages for example ActionScript, Java or C++. Some of these platforms are too complex to be useful to designers who would have basic coding knowledge at best. Ehn & Kyng (1991) and Rudd *et al*. (1996) warn that many computer supported prototypes rely on people in the design team being skilled programmers. Therefore construction of the interactive prototype might not be available to those that are involved in the design of the interface causing a potential for design intent to be 'lost in translation' between the designer and programmer.

2.4.3 Fidelity

When creating a prototype, the designer needs to balance the aesthetic and functional needs of the prototype, the environment within which it needs to operate (for example, user trials, demonstration, or a walk through), and the skills and resources of the prototype team (time and equipment). This balance of needs will have significant impact on the fidelity of the prototype.

Virzi *et al.* (1989) describe fidelity as being "a measure of how authentic or realistic a prototype appears to the user when it is compared to the actual service". Rudd *et al.* (1996) characterize low-fidelity prototypes as "limited function, limited interaction prototyping efforts *…+ constructed for illustrating concepts, design alternatives and screen layouts". The authors continue by defining high-fidelity prototypes as being "fully interactive" meaning that a user can "interact with the user interface as though it is a real product". Nilsson & Siponen (2006) proposed that fidelity can be explained by how automatic the response production is from non-automatic (facilitator driven) to fully automatic (user-driven).

This concept of 'low fidelity' prototyping began in 1990 with authors such as Nielsen (1993, p. 93), Virzi (1989), and Tullis (1985). Prototypes need to reflect the stage that the design concept has reached, there is no point producing a highly authentic looking prototype if the function of the product is still unknown. *"Prototype early and often, making each iterative step a little more realistic."* (Moggridge, 2007, p. 643). The need for more complex prototypes will become apparent as the design develops and more is known about the concept. Römer *et al*. (2001) sought information on how external representations were used in industry by conducting a detailed survey with practicing product and engineering designers. Simple models (low fidelity) were found to be of greater importance in the concept stages of the design process, with complex ones (high fidelity) only being created when more detail has been specified. Rudd *et al*. (1996) agree, stating that both low and high-fidelity prototypes have a place in the design process.

The appropriateness of fidelity, or the resolution, of a prototype is closely linked to the different user studies in which prototypes can be used. In the initial, exploratory phase of the design process, prototypes are made for two reasons: to explore ideas and to validate ideas (Broek, Horváth, & Lennings, 2000) (Lim, Stolterman, & Tenenberg, 2008).

A primary strength of a low-fidelity prototype is its incompleteness (Lim, Stolterman, & Tenenberg, 2008). The prototype needs to convey or explore what the designer needs, yet be so low-investment that it can be discarded once it has performed its function. The concept of fidelity is used to express this 'investment' in the prototype, with low fidelity being a lower resolution of the model in relation to the final design intent.

Virzi *et al.* (1996), found that there was little difference in usability data for high and low fidelity models of standard two dimensional graphical interfaces and an interactive voice response system concluding that the "*use of low-fidelity prototypes can be effective throughout the product development cycle, not just the initial stages of design*". Sefelin *et al*. (2003) demonstrated that a paper prototype produces equivalent results to a computer prototype but add that the participants preferred computer prototypes.

Lim *et al*. (2006) specifically aimed to prove the validity of using low fidelity prototypes in user testing. They pointed out that most comparisons to data had been anecdotal experiences or assumptions. They therefore constructed three prototypes of varying fidelity for user testing. The prototypes were based on a mobile phone, the high fidelity version was the real product, the medium fidelity was a screen-based prototype and the low fidelity version was a paper prototype. They found that all three prototypes identified the major usability issues, whilst the two 'real-time' interactive prototypes identified additional issues that the paper prototype did not. One of the identified limitations was that the graphical representation of the three-dimensional model resulted in a significant difference in 'physical manipulation and operation experience'. These prototypes demonstrate the value of low-fidelity techniques but there is a large gap between the final physical design and the two-dimensional paper prototyping and screen based prototype.

A number of researchers including McCurdy *et al*. (2006) and Lim *et al*. (2008) felt that the concept of low verses high fidelity is not quite enough to convey the whole manner of situations that prototypes are constructed for.

McCurdy *et al.* (2006) recognized that prototypes can have different strengths; they consider both a prototype that consists of only one or two non-interactive screens but is visually accurate to the pixel, and a prototype that uses the same input data and similar back-end logic as the delivered application, but whose visual look-and-feel is intentionally kept 'low resolution'. These are two very different prototypes that are very low fidelity in some respects but higher fidelity in others. In their paper, an approach to fidelity was proposed with five 'dimensions' of fidelity that can be defined as somewhere between high and low within the same prototype, namely, aesthetics, depth of functionality, breadth of functionality, richness of data and richness of interactivity. They trialled their five dimensions of fidelity approach on bespoke software which had three iterations: the 'current version' that is in use, a 'prototype' that consists of design changes from the current version and the developed 'proposed application'. The mixed fidelity prototype was low in two dimensions of fidelity (richness of data and level of interactivity) and high in the remaining three while the proposed application was high in all. The current version provided a baseline against which data from both prototypes could be compared. The proposed application sought to improve on a specific task identified in the current version. These tasks formed the basis of a usability study and they found that the mixed fidelity version was able to accurately predict the performance for the proposed application and that these both outperformed the current version. So far, this concept of mixed fidelity has been trialled with software but not physical prototypes.

Lim *et al*. (2008) classify prototypes by their features and not how they are used: they propose an 'anatomy of prototypes' accommodating the two key aspects of prototypes; as filters and as manifestations of design ideas. 'Prototypes as filters' recognises that prototypes are incomplete and are therefore intended to generate and evaluate ideas. They define five 'filters' for prototypes of interactive systems: appearance, data, functionality, interactivity and structure (relationships within the product). 'Prototypes as manifestations of design ideas' recognises that prototypes are (only) representations of an idea; the prototype is used to develop an idea and is not the output of the design. Manifestations of design ideas need to be considered in terms of the materials used to construct it, plus the resolution of the prototype and how 'complete' it should be; its scope.

Liu & Khooshabeh (2003) separated automation from fidelity and studied the effects of varying each with a Ubicomp software application prototype. They compared a paper prototype to an on-screen interactive prototype, and found that paper prototyping is insufficient for supporting the level of automation required, but a prototype with higher fidelity and automation levels can enhance the quality of interaction data available for evaluation. Sefelin *et al*. (2003) compared paper prototyping and computer-based prototypes of software interfaces and found that paper and computer-based low-fidelity prototypes lead to almost the same quantity and quality of critical user statements, but that participants prefer computer prototypes.

Sauer *et al*. (2008) constructed a table of 10 studies detailing the level of fidelity, type of model produced (two-dimensional or three-dimensional), the outcomes and the findings. They identified that most 2D prototypes had modelled 2D products rather than 3D products, and raised the question "to what extent would a 2D paper prototype be suitable to predict user behaviour for a fully operational 3D product". Thus, they focussed their research on the labelling of controls on both a paper prototype and a functioning pressure washer. They found that the "real appliance provided a somewhat different picture of user behaviour than the paper prototype". As part of their conclusions they propose that choosing a prototype of a certain fidelity level places constraints on the selection of task scenarios, the social and physical environment, and the user. Therefore, fidelity should not just be regarded in terms of the prototypes but also the wider testing environment.

Table 1 (page 58) provides an overview of current literature concerning fidelity levels in a format similar to Sauer *et al*. (2008). It shows the type of product being examined in relation to fidelity levels and a description of the construction of the prototype (detailing physicality at the simplistic level of two or three-dimensional). Some of the key studies which have developed knowledge of fidelity have been included even if they are software based. The focus was on physical prototypes for which the criteria for inclusion were: a) that the intended design is a physical product (as opposed to software), and, b) more than one prototype of the same design intent was created at varying fidelity levels.

This table highlights three key insights; 1) how the difference between low and medium fidelity is made, 2) how the physical product is represented, and 3) the types of product under study.

1. How the difference between low and medium fidelity is made

The distinction between a low and a medium fidelity prototype by Sauer *et al*. appears to be based on the level of automation with low fidelity being paper or cardboard-based and medium fidelity being computer-based. Yet the studies presented here demonstrate that basing the level of fidelity on automation alone is perhaps oversimplifying the matter.

2. How the physical product is represented

Interactive prototypes have traditionally been referred to by their fidelity, yet research into fidelity has focused predominantly on software only prototypes (McCurdy, Connors, Pyrzak, Kanefsky, & Vera, 2006), (Sefelin, Tscheligi, & Giller, 2003) (Liu & Khooshabeh, 2003). The table shows that of the research that does focus on physical interactive prototypes, the construction of the physical prototype is rarely typical of the product design process. For example; Lim *et al.* (2006) use a real mobile phone and vary the level of fidelity of the on-screen interaction; Virzi *et al.* (1996) use a paper keyboard but not a physical model for their electronic book and Sauer *et al.* (2010) overlaid a cardboard mock-up over the real appliance. In addition, McCurdy *et al*.'s (2006) proposal of the five dimensions of fidelity relates largely to the digital elements of the prototype and Lim *et al*.'s (2008) mixed fidelity approach can be applied without a physical model being present.

In a majority of studies identified during the literature review, the physical model was either an adapted final product or a very rough cardboard model. The study by Lim *et al*. (2006) was the only study to create a 3D foam prototype typical to the product design process but this was constructed to 1.5 times the size of the intended design in order to accommodate the paper screens. In no study was there a low fidelity model constructed with rapid prototyping technologies. Studies which specifically investigated the more 'complex' applications such as mobile phones all used final products for physical prototypes.

3. The types of product under study

For those studies that did implement an effective three-dimensional low or medium fidelity prototype, the product under trial generally had simple interactivity. For example, the can recycler (Sade, Nieminen, & Riihiaho, 1998) had three possibilities on insertion of the can (accept with refund, accept without refund and reject) and the pressure washer (Sauer, Franke, & Ruettinger, 2008) had two controls (pressure and temperature).

Table 1: An overview of published work on fidelity

Fidelity of prototype			Findings of study (*Critique*)
Low[a]	Med[b]	High[c]	

(Virzi, Sokolov, & Karis, 1996)	Study 1: Portable electronic-book player running an abridged encyclopaedia		
	Study 2: Interactive voice response system		
2D Simulation of the screens and keyboard on paper.	-	3D Final device.	Low-fidelity prototypes are as effective as high-fidelity prototypes at detecting usability problems. *(This was more akin to experience prototyping; physical models used were of existing products.)*
3D Wizard of Oz interaction on a real telephone.	-	3D High fidelity software implementation.	

(Sade, Nieminen, & Riihiaho, 1998)	Two concepts of a drinks can refund machine (manual and automatic)		
3D Both prototypes were constructed in 3D using foam board. Interaction was through the WoOz (with the Wizard in sight) who processed the cans and a facilitator changed the 3 LED lights.	-	3D Fully functioning prototype. LED functionality was not included.	No difference was found between 3D paper prototype and actual product. *(Simple product with no complex menu structure.)*

(Hall, 1999)	Domestic Lighting controller		
3D Cardboard mock up with images of the pushbuttons. Cardboard insert to simulate the LED's. Studied an original and a revised design.	2D Touch screen interface -lacked realistic tactile feedback of the push buttons, sound was added. (This was a student project, therefore many hours coding). Recreated both the original and the revised design.	-	The interactive 2D touch screen prototype revealed more usability problems than the 3D cardboard prototype. However the authors pointed out that the major usability issues were found by both prototypes. Additional issues were found because of the addition of sound in the touch screen prototype. *(Although extra functionality was not intended, the inclusion of auditory feedback in the touch screen prototype could be considered functionality as opposed to extra interactivity.* *Considerably more time was spent on the computer-based prototype than the paper prototype.)*

(Rooden, 1999)		Blood pressure monitor for home use	
2D Paper prototype.	3D Non-functioning final product.	3D Final product.	The usability errors identified with the real product were more similar to the usability errors identified with 2D paper prototype than with the 3D mock-up. *(No clear description of how the mock-up functioned. The high fidelity prototype was a final product.)*

(Sefelin, Tscheligi, & Giller, 2003)		Study 1: Calendar system	
		Study 2: Touch screen ticket system	
2D Hand-drawn paper prototype.	2D Computer-based prototype (same functionality as the paper prototype).	-	This study showed two main results: (1) paper and computer-based low fidelity prototypes lead to almost the same quantity and quality of critical user statements and (2) subjects prefer computer prototypes. *(These were both largely software orientated.)*
2D Hand-drawn paper prototype.	2D Computer-based prototype (same functionality as the paper prototype).	-	

(Liu & Khooshabeh, 2003)		Ubicomp (ubiquitous computing) software application - Kitchen-Net supports the task of working in an industrial kitchen by responding to spoken queries for items	
2D Paper prototyping: paper sketches and Post-It notes to represent Kitchen-Net screens.	2D Computer prototype - using Handheld PCs. This prototype automatically responded to events logged by the wizard, removing the need for the human 'computer'.	-	Both prototypes captured the major usability issue, the pc based prototype captured an extra two issues. *(The authors separated fidelity from automation, however many authors describe automation as a part of fidelity.)*

(Lim, Pangam, Periyasam, & Aneja, 2006)		Mobile phone	
3D Paper prototype interface operated on a foam prototype but 1.5 times the intended size. (5 hours construction time for 3 people).	2D Laptop–based screen emulation (20 hours construction time for 2 people).	3D Final device ("because it most precisely represents the actual design").	Major usability issues were found by all prototypes. Further issues were uncovered by the medium OR high fidelity prototypes. *(The low fidelity prototype used an existing phone case, therefore its physicality was not low fidelity.)*

(McCurdy, Connors, Pyrzak, Kanefsky, & Vera, 2006)	Activity-planning tools for Mars surface operations		
-	2D On screen prototype.	2D Two final software programs (one existing and one proposed).	No significant difference between the two prototypes. *(Seminal study but only trialled on purely software prototypes.)*

(Sauer, Franke, & Ruettinger, 2008)	Pressure washer		
2D Paper prototype, a colour photograph of both pressure and temperature control plus photographed scenarios.	-	3D Final physical device with modified control labels.	The results showed that the real appliance provided a somewhat different picture of user behaviour than the paper prototype, suggesting the need for caution in interpreting behavioural data obtained with a paper prototype. *(This study focused on eliciting feedback specifically concerning the graphical labelling of controls.)*

(Blackler, 2009)	Microwave		
2D Paper prototype (on a vertical wall).	2D PowerPoint interface on a touchscreen (approximately half of intended size). A cardboard 3D model was used to introduce participants to the intended size and physicality.	-	The authors found that both types of prototyping have a role in the design process: this study shows that high-fidelity touchscreen prototypes can be successfully employed as experimental tools and that low-fidelity prototypes may have an important application in developing design tools. *(Despite a non-functioning prototype realised in 3D, this is a largely 2D interface.)*

(Sauer & Sonderegger, 2009)	Mobile phone		
2D Paper prototype - sketch representation of physical model.	2D Computer-based simulation (PowerPoint) interacted through a touch screen (ThinkPad).	3D Final devices (Sony Ericsson and Motorola).	The main results showed that task completion time may be overestimated when a computer-based simulation is used. The effects of fidelity levels on attractiveness ratings appeared to be stronger for less appealing products than for attractive ones. Objective performance parameters collected during the usability test and subjective usability ratings were not associated. Results showed no evidence for the fidelity level affecting emotions or subjective user evaluation. *(Only 2D prototypes were used)*

(Sauer, Seibel, & Ruttinger, 2010)	Floor scrubber		
3D 2D representation of the interface containing all controls in a simplified form with regard to the aesthetic refinement and tactile representation. The paper prototype was modelled in cardboard (sized 300mm 300 mm), upon which all possible configurations were drawn. The controls were made of foam rubber that was fixed by a paper clip on the cardboard allowing their pushing or turning.	3D Partially operational 3D mock-up with all navigational functions fully available (e.g. speed). PVC/cardboard mock-up over the real appliance. Mock-up had duplicates of the all functions (e.g. water flow rate) but the functions were non-operational so that they had no effect on cleanness levels. Since the real appliance was completely covered with the mock-up, users had the impression that they were operating a not fully operational prototype.	-	Reduced fidelity prototypes for determining user behaviour with real appliances may lead to a general overestimate of control settings since users employing a reduced fidelity prototype chose generally higher control settings than those using the real appliance. *(The medium fidelity level prototype used the final product as a basis -what would the authors have done if this was not available?)*

From Sauer et al (2008): [a]paper or cardboard [b]computer-based [c]final product or high fidelity prototype

2.4.4 Summary

Prototypes are physical manifestations of the design process. Interactive prototypes are responsive to a user's action, either directly in the case of 'smart' prototypes with embedded electronics, or appear to respond through 'faking' the interaction.

Interactive prototypes have had considerable attention from researchers with a lot of focus on fidelity. However, most of these researchers have focussed purely on the digital aspects of the prototype or, at best, given the physical prototype limited consideration. An opportunity exists to specifically explore the usefulness of the physical aspects of the interactive prototype.

The Contextual Study presented in the next section develops the understanding of physicality in relation to computer embedded products through a user trial on a series of low-fidelity prototypes.

2.4.5 Contextual study -'Mobile Home Phone Study'

This study sought to explore the effect of fidelity on prototypes of handheld interactive products, the full paper can be found in Appendix 5.

> Gill, Walker, Loudon, Dix, Woolley, Ramduny-Ellis, & Hare (2008), **Rapid Development of Tangible Interactive Appliances: Achieving the Fidelity/Time Balance**. In Hornecker, E., Schmidt, A., and Ullmer, B. (eds) Special Issue on Tangible and Embedded Interaction, International Journal of Arts and Technology, Volume 1, No 3/4 pp 309-331.

2.4.5.1 *Overview of the study*

The objective of the study was to determine if the results of a user trial with a tangible prototype were more similar to the final product than a software-only prototype, and the subsequent level of fidelity required of this prototype. The results suggested that it is not the level of fidelity that is important but rather the considerations of tangibility and physicality.

2.4.5.2 *The prototypes*

Four prototypes were constructed; the first two being a high-fidelity model and a software-only prototype (mimicking common prototyping practices), and a further two that lowered the level of fidelity of the prototype, these are shown in Figure 16.

Prototype 1 (named 'high-fidelity') was created by connecting an IE Unit (a modified keyboard chip) to buttons in the casing of the final product, the IE Unit enabled button presses on the phone to trigger a mock-up of the phone's interface created in Flash and shown on a laptop. The same Flash interface was used for **Prototype 2** (named 'software-only') and operated through a touchscreen laptop. **Prototype 3** (named 'sketch') consisted of a blue foam model of the product with basic integrated buttons and reduced functionality sketch graphics within the Flash interface. The IE Unit was again used to connect the physical model to the computer to operate the interface. For **Prototype 4** (named 'flat-face') a blue foam model was created, but, in an effort to reduce the time taken to construct the prototype, the buttons were not embedded into the front of the completed foam model. Instead the front face of the foam model was left flat while the back was modelled as normal; the buttons were then embedded into the flat face of the model. A paper print out was used to cover the physical buttons creating an impression of the final product. The same sketch Flash interface was used as for prototype 3.

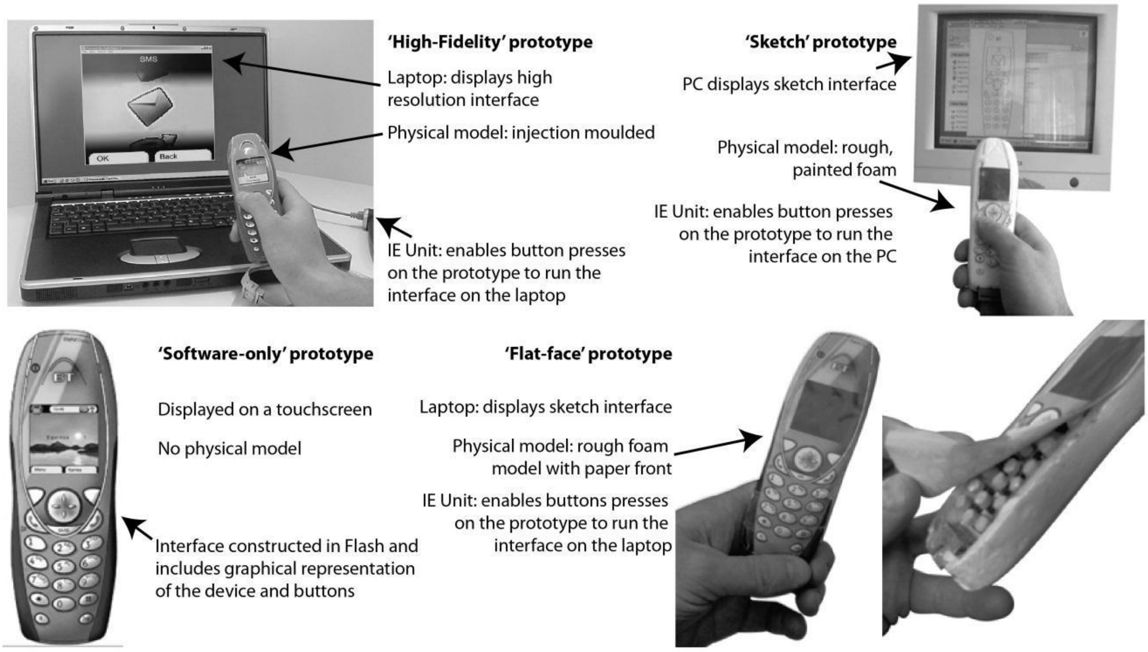


Figure 16: The four prototypes of the Mobile Home Phone Study

2.4.5.3 ***Methodology***

A between-subjects user trial was conducted with 16 participants for each of the four prototypes plus the real product. The participants were asked to complete six tasks and the success rate of each task (Molich & Dumas, 2008)) was recorded along with the time taken to complete it.

2.4.5.4 ***Results***

The 'high-fidelity' prototype produced similar results to the final product, significantly outperforming the 'software-only' prototype. The 'sketch' prototype was found to perform similarly to the final product. The performance of the 'flat-face' prototype, however, was significantly reduced. It appeared that the flat face of the prototype did not replicate the true physicality of the product with sufficient accuracy, this resulted in a high rate of user errors which produced slower performance times and poor performance ratings. The results suggest that it is not the level of fidelity that is important but rather the considerations of tangibility and physicality.

2.4.6 Conclusion

This theme focused on the prototype as a physical expression of the design process. The literature review identified the concept of fidelity as fundamental to understanding the use of prototypes. However, the literature review also identified a lack of focus on the physical aspects of the prototype. Most research papers agreed that low-fidelity prototyping methods were very effective in comparison to higher fidelity techniques. Yet many of those studies focused largely, if not entirely, on the digital elements of the design, forsaking the physical.

The study of the 'Home Phone' revealed the importance of the physical prototype, especially for the computer embedded product where the design of the interface and the physical product is often separated. The indication that the physicality, not fidelity, of a prototype can affect the results of user trials proved a valuable insight. This study suggested that, with a better understating of physicality, low-fidelity prototypes can produce good quality data in user trials of computer embedded products combining both physical and digital elements.

2.5 Conclusion and Development of the Research Question

The focus of this PhD is *physicality in relation to the design and development of computer embedded products*. The literature review and contextual studies explored three themes relating to physicality in the design of computer embedded products; the meaning of physicality in relation to computer embedded products, the design process of computer embedded products, and the physical manifestations of the design process (specifically interactive prototypes).

Physicality can be understood to be the physical aspects or qualities of both an object and an interaction; this includes our physical bodies in relation to that object.

Physicality is central to our experience of computer embedded products, from how we exist in our bodies within the physical world, and how we perceive interactions with the physical world to the point at which we interact with that physical world.

Each of the three themes explored various aspects of physicality and highlighted important areas to be considered by the research question.

Theme 1 identified the importance of addressing the affordances of the prototype and the potential for considering the prototype 'unplugged'.

Theme 2 identified that physicality is intrinsically part of the way computer embedded products are designed, therefore identifying the need to address the way in which designs are realised.

Theme 3 indicated that physicality could provide the basis for creating prototypes which can produce good quality data in user trials. Investigating this theme calls on the knowledge gained in Themes 1 and 2 whilst providing a focussed investigation of the influence of physicality.

In light of the review of literature and the studies presented here, the following research question was formulated:

> **Can a better understanding of physicality help in the creation of more effective low-fidelity physical interactive prototypes?**

Chapter 3. Methodology

The research question addressed by this thesis is:

> Can a better understanding of physicality help in the creation of more effective low-fidelity physical interactive prototypes?

As such, the methodological framework is based on the evaluation of multiple prototypes of the same design intent in two stages; first a comparative analysis of the physicality of the prototypes and second, a comparative analysis of the prototypes through user trials.

The data for this thesis was gathered through two studies; Study One investigated the relationship between fidelity and physicality, and Study Two investigated the direct effect of physicality on low fidelity prototypes.

The design of the studies was based on two independent variables; the design intent of the product and the structure of the trials. In each study, the prototypes were constructed with the same design intent including the same functions and features. Each prototype within the study was tested with an identically structured user trial.

In addition, each case study had a specific independent variable which determined the method of prototype construction, these were; time limitations for Study One, and physicality levels for Study Two. The resultant prototypes were dependent on these parameters and the impact on physicality could then be assessed. Once physicality had been analysed and the prototypes have been trialled with a consistent structure, any differences in the results of the user trials can be compared in relation to the physicality of the prototype.

The next sections detail the key considerations of the methodology; the generation of the prototypes, the comparative analysis of the physicality of the prototypes, the comparative analysis of the prototypes through user trials and ethical considerations of the research.

3.1 Generation of the prototypes

Multiple prototypes were created for each study to enable comparison across the prototypes. There are two main considerations for the generation of the prototypes; what to prototype and how to prototype it.

3.1.1 What to prototype: retro-prototyping versus a conceptual design

The choice between retro-prototyping an existing product and creating a conceptual product was given considerable thought. Both methods have been used in prototype evaluation methods found in the literature review (as shown in Table 1 on page 58). The benefit of retro-prototyping is that the existing product simply needs to be reproduced at prototype level; there is no design work to be undertaken resulting in a faster study design. There is the added benefit in that a real product exists that can be used to benchmark the prototypes, indeed this was the approach taken in the study of the home phone by Gill *et al*. (2008) and presented in the contextual studies of 2.4.5. The major compromise of this approach is that all the design decisions would have been made; consequently any prototype constructed will not fully reflect typical early stage prototypes when there are many unresolved aspects of the design.

As such, for the first user trial, the decision was made to produce a new product concept in order to ensure early-stage testing was done with real early-stage prototypes, and not a reverse engineered end product of an unknown process. On reflection of Study One, this approach resulted in difficulties in the analysis when comparing results from across the prototypes without a benchmark to compare against, therefore an existing product was retro-prototyped in Study Two.

3.1.2 How to prototype it: construction techniques and limitations

The difference between the prototypes results from the choice of construction technique. The literature review highlights a number of tools and techniques for creating low-fidelity prototypes, including; Phidgets (Greenberg & Fitchett, 2001), the IE System (Gill, 2013), Arduino (Burleson, Jensen, Raaschou, & Frohold, 2007), DTools (Hartmann, Klemmer, Bernstein, & Mehta, 2005) and Wizard of Oz (Buxton, 2007).

For each study, the decision of how to construct the prototype was determined by the study design; Study One was constrained by time in order to explore its effects on fidelity and the subsequent physicality, and for Study Two, the development of the construct of physicality enabled a more focused investigation of physicality as opposed to fidelity.

3.2 Comparative Analysis of the Physicality of the Prototypes

To enable a comparative analysis of the prototypes, the differences in physicality between the prototypes must first be understood. Unfortunately there is no precedent in assessing the physicality of a prototype. The framework by which to assess physicality has therefore formed part of the main strand of this PhD and the framework has evolved accordingly.

Initially, a comparable approach to understanding the qualities of a prototype was sought. Research into the level of fidelity of prototypes, as introduced in the literature review (Section 2.4.3 on page 53), has provided a starting point into this investigation into physicality. Therefore, the study of fidelity seems a logical starting point from which to determine a means of comparing prototypes.

A number of different models of fidelity are covered in the literature review. These include; the widely used low versus high fidelity model (Nielsen J. , 1993) (Virzi R. A., 1989); non-automatic versus fully automatic (Nilsson & Siponen, 2006); the 'Five Dimensions of Fidelity' proposed by McCurdy *et al*. (2006); and the 'Anatomy of Prototypes' by Lim *et al*. (2008).

These models recognise the intricacies of the interactive prototype by reflecting the various facets that comprise the completed interactive prototype. In each of the models proposed, the prototype is broken into facets which can be assessed individually. The limitation of some of these models is that the 'product' is often described as a website or piece of software and neither of these is the focus of this research. That said, there certainly seems to be the opportunity to bring across some of the ideas that have been proposed through the study of the fidelity of software products into the study of computer embedded products.

The 'Five Dimensions of Fidelity' model by McCurdy et al. (2006) was used as a starting point to compare the prototypes of Study One. Firstly, the 'Five Dimensions of Fidelity' were used to assess the fidelity and subsequently, the model was adapted to make it more relevant to physicality. Instead of five dimensions, two 'areas' of physicality were identified, these were: physicality in relation to the product and physicality in relation to the interaction. Within these areas 'driving factors' were identified which had a direct effect on each area of physicality. These drivers were specific to the prototype being constructed and not generalised terms, for example, how the dial mechanism felt.

For Study Two this was completely revised by the proposal of active and passive physicality (as explained in Section 4.11).

3.3 Comparative Analysis of the Prototypes through User Trials

In order to determine the effects of physicality on the prototypes, a way of establishing the 'effectiveness' of the prototype needs to be determined. The literature review identified user trials as a common reason for creating prototypes. In this research, an 'effective' prototype is defined as one that elicits meaningful comments and insights during user trials conducted during the early stages of the design process. Here, 'meaningful' comments are those that focus on improving the overall intended design of the concept as opposed to the interface in isolation or the construction technique of the prototype.

User trials are simulations of product usage in which subjects are asked to fulfil specified tasks using a product or prototype (Vermeeren, 1999). User research is an effective way of soliciting feedback from the target user group, the results of which can be fed back directly into the design process. As Nielsen (1993, p. 165) describes, a user trial with real users *"is the most fundamental usability method and is in some sense irreplaceable, since it provides direct information about how people use computers and what their exact problems are".*

User trials have been used by many researchers who have investigated the effect of fidelity on a series of prototypes, including Sauer et al. (2010) and (2008), McCurdy et al. (2006), Lim et al. (2006) and Virzi et al. (1996). The investigation of the effects of physicality, which is the focus of this thesis, compares multiple prototypes, and user trials can therefore be used with some degree of confidence. User trials provide empirical data within a subjective topic that enables comparison across prototypes. The aim of comparing the prototypes through user trials was to gather data that enabled a review of the differences in the *way* the prototypes function as each of the prototypes has the same *level* of functionality.

3.3.1 The User Trials

The design of the two user trials presented in this thesis was based on the same underlying structure, although there were some differences which resulted from the aim of each study. Study One was open and exploratory, and reflected the structure of the Home Phone contextual study presented in Section 2.4.5. Study Two was more focused, and lessons learnt from Study One were applied to Study Two. In addition, facilities became available in Study Two which enabled the trials to be held within a user lab with dedicated cameras.

There are several considerations that were common across both user trials; the next sections describe these considerations.

3.3.1.1 ***Location***

The trials need to be held in a location that can accommodate the necessary equipment and people. Many user trials are undertaken in a dedicated space, but this does not necessarily reflect the context within which the final product will be used. Kjeldskov and Graham (2003) found 'a clear bias towards environment independent and artificial setting research' for mobile computer embedded products, however, Woolley (2008) found that testing in context uncovered physical problems not seen in the laboratory setting.

In a lab-based trial, participants are bought into an area that has been pre-configured for the user trial for which the primary goal is to provide a consistent, quiet, comfortable space to do research (Kuniavsky, 2003). This could be as simple as setting up a camcorder on a tri-pod, or an entire dedicated environment, and indeed both configurations have been used in this PhD research.

User Study One was conducted in an 'ad hoc' space with a camcorder positioned to record the prototype in the participants' hands. Between the studies, PDR commissioned a new dedicated user lab and User Study Two was held within this Insight Lab.

The Insight Lab at PDR has four cameras that are controlled from a separate observation room. The cameras can pan and zoom to enable a scene to be captured from the most appropriate angle or angles. There are three audio points within the room to capture conversation.

There are strengths and weaknesses to both configurations. The 'ad hoc' lab required no extra people to operate the equipment, plus it could be moved easily to a different location to make it easier to take the 'lab' to the participants. The weakness of this configuration is that the camera is very obvious to the participant, and it required monitoring to ensure it captured the appropriate data. In addition, the ability to move the location of the lab could introduce inconsistencies outside the control of the moderator, such as noises or the movement of others through a space.

The dedicated lab kept the environment consistent and the cameras less intrusive. The weakness of this configuration was that each participant had to come on site and an additional person was required to operate the cameras, both making sure they were angled appropriately and by ensuring the video feed (or feeds) were being recorded at the correct time.

3.3.1.2 ***Configuration***

At its very simplest a user trial needs two people, a **participant** who has been identified as a potential user, and the **moderator** who will ask the participant questions and set 'tasks' to perform on the product.

Depending on the configuration of the environment, further people might be required. If the cameras can be controlled remotely, an **operator** is required. Study two was conducted in PDR's Insight Lab which has four independent cameras controlled from an observation room. In this set-up an operator is required to start the cameras recording once the participant's consent form has been completed.

Another role could be a **facilitator** to operate the prototypes; this role is required if the prototypes need setting up between trials or tasks, or if there is an external trigger event during the study. In the studies presented here, the moderator also performed the role of the facilitator by operating the prototypes. Although care needs to be taken using this approach because it can introduce a bias if the participant knows the moderator created the item under trial (Nielsen J. , 1993, p. 188). This bias is created because participants are more likely to try and 'please' the moderator if they think they are judging the moderator's work. For both the studies here, it was made clear that the prototypes were constructed elsewhere and the participants were asked to be as truthful as possible.

Any form of user trial will affect the user, when people know they are being observed, their behaviour changes. This is known as the Hawthorne or Observer Effect (Hart, 1943). Steps can be taken to reduce the observer effect such as: keeping the number of people involved to a minimum (at a visual level at least), keeping recording equipment unobtrusive and using an experienced moderator who can make the participant feel comfortable in expressing their thoughts.

3.3.1.3 *Within-Subjects versus Between-Subjects Study Design*

The common aim for both studies was to compare a variety of prototypes of the same functionality. There are two ways in which multiple designs can be compared; by asking a user to use all the designs (within-subjects), or asking a user to use one and then compare multiple users (between-subjects). For a within-subjects study design a participant will use all versions of the prototype enabling the participant to give feedback on all prototypes, applying their individual knowledge in each case. The danger of this approach is that the participant will learn how to conduct a task having been asked to perform it on the first prototype, and use that new knowledge to help complete the task on subsequent prototypes. For a between-subjects study design, each participant will only perform the tasks on one prototype so there can be no effects of learning; however this means a larger number of participants need to be recruited for each prototype.

Tullis and Albert (2008) address the need to compare alternative designs in their book. They recognise that a within-subjects design would not work due to the learning effects resulting from similarity of the designs. Two options are suggested, a purely between-subjects study design, which requires a greater number of participants, or a mixed approach where tasks are performed on alternating designs with the remaining designs introduced to ask for a preference.

All the prototypes under consideration for this PhD have the same functionality, and therefore tasks can be solved in the same way for each prototype, and the potential learning effect between each prototype is considerable. Therefore, a between-subjects approach was adopted for both studies. Study One was purely between-subjects but it was realised that an opportunity was lost for participants to give their thoughts on the other prototypes. The inclusion of such an opportunity after the task-based user study would provide a within-subjects study of all the prototypes. Therefore, this mixed approached was adopted for Study Two where users first undertook between-subjects trials on one of the prototypes, and were subsequently introduced to all the prototypes using a within-subjects questionnaire.

3.3.1.4 ***Number of Participants***

The number of participants to engage in a user trial is of some debate; Tullis and Albert (2008) provide an excellent overview of the discussion on pages 117 – 121 of their book 'Measuring the User Experience'. Tullis and Albert state that the two sides of the argument are based on whether five users are enough to provide a majority of the feedback. Nielsen and Landauer (1993) devised a graph (Figure 17) which demonstrates the number of usability issues identified after each participant; the graph shows that 80% of usability issues are uncovered by 5 participants. Nielsen and Landauer argue that time is better spent conducting a series of iterative tests rather than more users in a single iteration.

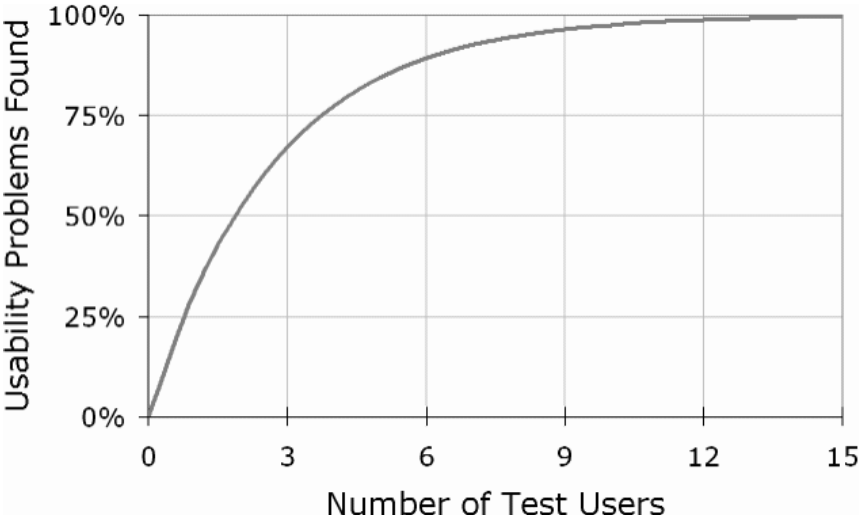


Figure 17: Nielsen and Landauer

Others, including Molich *et al*. (1998) and Spool & Schroeder (2001), have suggested that over 10 participants are needed to account for individual differences. Tullis and Albert conclude saying that during the early stages of design, fewer participants are needed to identify the major usability problems and five participants provide enough data **IF:**

- the scope of the evaluation is limited to 5-10 tasks,
- **AND** the participant screening is an effective representation of the intended user group.

As noted earlier, the study design needed to be between-subjects. Therefore, 16 participants were recruited for each of the three prototypes for Study One, resulting in 48 participants in total. Study Two required trials on four prototypes plus the final product, therefore 8 participants were recruited per prototype (including the final product), resulting in 40 participants in total. In addition, for Study Two, all participants were shown all the prototypes at the conclusion of the tasks. This allowed data to be gathered from all 40 participants in relation to all five of the prototypes.

3.3.1.5 ***Structure***

There are several resources that provide guidance of how to conduct a user trial. The books 'Measuring the User Experience' (Tullis & Albert, 2008), 'Observing the User Experience' (Kuniavsky, 2003) and the 'Handbook of Usability Testing' (Rubin, 1994) provided the basis for the design of these studies.

The general structure was as follows:

1. Meet and greet –*the participant is met and shown to a waiting area.*
2. Participant information sheet –*while waiting the participant can read the information sheet and is given the opportunity to ask questions.*
3. Consent form –*once any questions have been answered the participant signs the consent form, from this point on participants are assigned a number to keep their data confidential.*
4. Demographic questionnaire –*opportunity for demographic data to be collected.*
5. Brought into lab if not there already –*introduced to the lab or location of the trial.*
6. Introduced to the product and structure of the session.
7. Perform the tasks.
8. Post-task questions –*capturing any information related directly to each task.*
9. Semi-structured interview –*a series of open questions concerning the product in general*
10. Any questions/debrief.
11. Finish

Both of the studies in this research had five tasks, one hour was allocated for each participant and in the event the sessions generally lasted 40 minutes.

3.3.1.6 ***Tasks***

The most common usability metric is ‘task success’ (Tullis & Albert, 2008). A task is a predefined goal representative of typical user activities and sufficiently isolated to focus attention on a single feature (Kuniavsky, 2003). The tasks needed to be defined early in the development of the prototypes in order for the appropriate functionality to be enabled across the prototypes. If the user strayed too far from the task a screen displaying ‘feature unavailable’ would be shown. Tasks typical of the product were chosen and implemented within all the prototypes.

There are many guidelines for constructing an appropriate sequence of tasks. Kuniavsky (2003) lists seven features of a ‘good’ task: a reasonable task, described in terms of the end goal, specific, do-able, performed in a realistic sequence, non-specialist for the potential users and of a reasonable length.

Tasks can be of varying complexity and ‘simple’ tasks were typically used at the beginning to make participants feel comfortable. Tasks were also in an order of ‘typical’ usage, i.e. the product needs to be turned on before it can be used and headphones need to be plugged in before listening to music. Each task must have a clear end-state and be specific, for example finding a specific piece of information.

3.3.1.7 ***Think aloud protocol***

The ‘think aloud protocol’ encourages participants to verbalise their thoughts as they are working through the tasks (Tullis & Albert, 2008, p. 103). Lewis and Rieman (1993) are credited with introducing the protocol to the usability field; they suggest that the moderator should ask the participants to talk about what they are thinking, such as what they are trying to do, questions that arise as they work and what they read. Encouraging participants to talk without leading them is very important, therefore Lewis and Rieman suggest the moderator use phrases such as “*tell me what you are thinking*” rather than asking “*why did you do that?*”.

3.3.1.8 ***Semi-structured interview***

An interview is a chance to directly ask the participants about their experience. A usability interview tries to completely remove the perspective of the person asking the questions from the interview (Kuniavsky, 2003). A semi-structured interview is based on some scripted questions that the interviewer can probe into as necessary. Questions should focus on a single topic and be open-ended. This approach ensures the consistency and repeatability of the interview whilst enabling freedom to explore a particular comment made by an individual participant.

3.3.1.9 ***Data captured***

The structure of the user trial presented allows both qualitative and quantitative data to be collected. The data was gathered to enable a comparison across the range of prototypes in a trial typical of early stage user trials which seek feedback concerning the overall design and basic information architecture.

Quantitative data

Performance metrics are used to determine how well users are actually using a product (Tullis & Albert, 2008). They can also be used to determine the magnitude of identified usability issues. One of the main metrics than can be recorded is the success rate for each individual task. The categories for rating success were based on those proposed by Molich & Dumas (2008) for reporting usability problems:

- Success: the participant completes the task without problems or delay.
- Minor problem: The participant is briefly delayed (the participant experiences a problem, but corrects themselves reasonably quickly (less than 1 minute)).
- Serious problem: The participant is significantly delayed (1 – 5 minutes) but manages to complete the task.
- Catastrophe: The participant is unable, or refuses, to complete the task or the participant solves the task incorrectly without noticing.

The moderator noted down the relevant performance metrics during each task, in addition, the usability problems that caused users to struggle with tasks were also recorded for later analysis. All ratings and usability problems were subsequently reviewed using video capture data to ensure consistency.

Qualitative data

The comments made by participants were obviously of importance for these studies, whilst the performance metrics focused on the usability of the products, the comments made by participants were expected to produce a rich insight to their opinion of the product as a whole. The semi-structured interview offered an additional chance to elicit participants' comments concerning specific aspects of the product. The entire study was recorded on video with audio and therefore comments made during the tasks and semi-structured interview were captured for later analysis.

3.3.1.10 ***Analysis***

Quantitative Analysis

Both user studies collected performance data during the tasks, based on whether the tasks performed were a success, minor problem, major problem or failure. This data was converted into interval data where 3 = success; 2 = minor problem; 1 = major problem and 0 = failure.

This interval data was then analysed to discover if there is a statistical difference between the prototypes performance using a mixed level of variance analysis (ANOVA). This type of analysis looks at the variation between three or more groups. In this case, the groups were the participants using each of the prototypes.

Data concerning completion time was gathered for Study One. This was averaged for each task with a confidence interval to incorporate variability.

Qualitative data

Extensive video data was collected during each study to capture the comments received. Discourse analysis (Jørgensen & Phillips, 2002) provided a framework to analyse the video data. This approach uses discourse such as speech or writing as a basis for analysis. The strength of this approach is that it gives the ability to structure the conversational feedback typical of this type of study in a more rigorous manner. Typically, formative user trials are not analysed to this level of detail, however this approach was taken due to the need to directly compare the data obtained for each prototype in a manner rigorous enough to draw insights for this PhD research.

The video footage from each of the participants was reviewed with event logging software and comments were assigned 'codes' based on the type of comment. Transana was used for user Study One and Observer XT for user Study Two. Transana is qualitative analysis software specifically for the analysis of digital video or audio data (Transana, n.d.). Observer XT is analysis software for the collection, analysis, and presentation of observational data (Noldus, n.d.).

Usability issues were then coded from the comments received and errors observed. A usability issue is the underlying cause of a problem which prevents the user interacting with the product in the intended manner.

Tullis and Albert (2008, p. 111) describe the most common way to measure usability issues is to count the number of unique occurrences. This is aggregated over each participant, as such, the second participant might encounter ten usability issues, but only five of those might be different from the first participant. Counting the usability issues in this way means that the later participants should each only uncover one or two new usability issues.

Alternatively, the 'frequency issues per participant' or the 'frequency of participants' (Tullis & Albert, 2008) can be recorded. The 'frequency issues per participant' is the number of issues each participant encounters, whereas the 'frequency of participants' is the number of participants who encounter a specific problem. An alternative way to analyse the usability issues is to categorise them, either by type of usability issue such as navigation or terminology, or by task.

In order to compare the prototypes, the usability issues were categorised by type for both the studies for this PhD research. The 'frequency of issues per participant' was captured for Study One, so that it would be recorded if a participant repeatedly encountered an error. These values could then be collated for each prototype to enable comparison of the 'frequency of issues'.

In Study Two, some of the participants were excessively vocal concerning a specific usability issue, if the 'frequency of issues' approach was taken from Study One these repeated issues were recorded, therefore the results did not fully reflect the study. To counter this, the data was compared by analysing the 'frequency of participants per usability issue'. In this approach, the usability issues were coded along with the number of participants who experienced that issue.

A similar approach was taken for both studies when analysing comments to determine if there were any trends which related specifically to physicality. For Study One, comments were categorised into three areas: the physicality of the product, the physicality of the interaction and feedback concerning the product in general. For Study Two, comments were assigned 'codes' based on the type of comment and 50 comment groups were recorded. These were refined into 10 design recommendations.

3.3.1.11 ***Limitations of user trials***

User trials are just one of many ways in which low-fidelity prototypes are used in the early stages of the design process. Other reasons for creating prototypes include; exploring the shape of a design, exploring how a design functions, demonstrating how the design might look or for stimulating discussion within a group. Ulrich & Eppinger (2012, p. 294) define four purposes of prototypes namely: learning, communication, integration and milestones.

User trials are recognised to have limitations; Snyder (2006) reviewed the ways in which usability findings might be biased, these have been distilled into six general categories by Tullis and Albert (2008, p. 116):

1) Participants –level of experience participants bring for example, how comfortable they are in the given situation.
2) Tasks –appropriateness of task chosen including wording of task used.
3) Method –way in which the study is evaluated.
4) Artefact –the nature of the prototype or product being evaluated has a huge impact on findings. The type of interaction will vary whether it is a paper prototype, functional or semi-function prototype or production system.
5) Environment –setting for the studies for example, outside influences or intimidating surroundings.
6) Moderators –experience of moderator, presumptions concerning study outcome (the evaluator effect).

The Comparative Usability Evaluation (CUE) studies, led by Rolf Molich, explored confidence in usability reporting. CUE-4 specifically compared the evaluation of the same interface (a website) by seventeen different professional teams (Molich & Dumas, 2008). This study found only 9 of a total of 340 usability issues detected were found by all seventeen teams. This indicates that different usability issues would be uncovered if a study were to be repeated by different teams. In all the studies for this research, the trials were designed, conducted and analysed by the author, giving confidence in comparisons across the prototypes. If each prototype had been compared by a different researcher, such confidence would not be possible. However this does introduce presumptions held by the author, specifically concerning physicality, which will inevitably influence the outcomes.

3.4 Ethical Considerations

Ethical approval was sought through Cardiff Metropolitan University's Ethics committee (formerly UWIC). Included in the documentation for ethics approval is the information sheet that is given to the participant informing them of what is expected of them if they agree to participate plus a consent form which they are required to sign if they agree to take part in the study.

3.5 Limitations of the Research Methodology

The prototypes could also be exclusively assessed in terms of the resultant physicality of prototypes created through the multitude of prototyping techniques available. This type of assessment could be expanded onto a variety of prototypes constructed for commercial design projects. The limitation of this route is that it would not give any indication of the effects of physicality on the 'effectiveness' of the prototypes because a range of prototypes would not be constructed or tested. That said, this could be a really insightful way to progress the findings of this research further.

3.6 Conclusion

This Chapter has discussed the way in which the research question was approached. The methodological framework is based on the evaluation of multiple prototypes of the same design intent in two stages; first a comparative analysis of the physicality of the prototypes and secondly, a comparative analysis of the prototypes through user trials.

Two studies were undertaken to address the research question: *Can a better understanding of physicality help in the creation of more effective low-fidelity physical interactive prototypes?*

Study One (Chapter 4), investigated the relationship between fidelity and physicality. Development work was undertaken after this study which proposed a potential way of 'better understanding' physicality through a framework of physicality. Study Two (Chapter 5) investigated the direct effect of physicality on low fidelity prototypes.

Chapter 4. Study One: a Conceptual Photo Management Product

4.1 Introduction

The user study of the Home Phone undertaken by Gill *et al.* (2008), and presented in the contextual studies in Chapter 3 (starting on page 62), concluded that it is not the level of fidelity that is important but rather the considerations of tangibility and physicality. Therefore, Study One investigated the relationship between fidelity and physicality by constructing a series of prototypes using time constraints to drive the fidelity and the subsequent physicality of the prototypes. User trials were then used to enable a comparative analysis. The hypothesis of this study was:

> *A relationship exists between fidelity and physicality and this relationship affects the effectiveness of the prototype.*

The following paper was published and presented at Interact 2009 and can be found in Appendix 6.

> Hare, Gill, Loudon, Ramduny-Ellis & Dix (2009), **Physical fidelity: Exploring the importance of physicality on Physical-Digital conceptual prototyping**. In the proceedings of Human-Computer Interaction - Interact 2009, Uppsala, Sweden. Springer Berlin Heidelberg.

4.2 Method

As discussed in the Methodology Chapter, a two-stage approach was taken to the research studies; the first stage was a comparative analysis of the physicality of the prototypes created and the second stage was a comparative analysis of the prototypes through user trials.

This two-stage approach sat within the wider study method which consisted of the following:

1) Choosing the product to trial
2) Generation of the prototypes
3) Comparative analysis of the physicality of the prototypes
4) Conduct the user trials
5) Comparative analysis of the prototypes through the user trials.

The next sections provide further detail for each of these steps.

4.3 The Product under Trial

As discussed in the methodology chapter, a conceptual product was chosen for Study One, this ensured early-stage user testing was done with 'real' early-stage prototypes, and not a reverse engineered end product of an unknown process. The conceptual product originated from an undergraduate design brief named the 'Flickr Friend'. Flickr is an online photo management and sharing application. The design used for this study was based on a hard drive equipped product that offers users the ability to wirelessly view their Flickr web pages and to store photos.

Figure 18: Physical design of the conceptual photo management product

The concept of the product is based around Flickr 'local' which refers to the storage of photographs 'locally' on the hard-drive, and Flickr 'live' which refers to the storage of photographs on the Flickr website. Once the product is turned on, users can choose between Flickr local and live. The users' photographs can be transferred from a camera to the product and viewed through Flickr local by year and by set, from here, photographs can be selected for upload to the Flickr website (Flickr live). On Flickr live the user can view their friends' photographs, as well as view and manage their own images.

Figure 18 shows the final design of the product; it has a large screen on the front with a dial on the back. There is a single button visible from the front and a further two buttons on the sides, one on the top and one on the base. Figure 19 shows the interactions with the product. The user first turns the product on with the button on the base (1). The wheel is used to scroll between menu items (2). The current menu item can be selected with the ‘S’ button on the front of the product (3). Finally, the ‘back’ button is located on the top edge of the product (4).

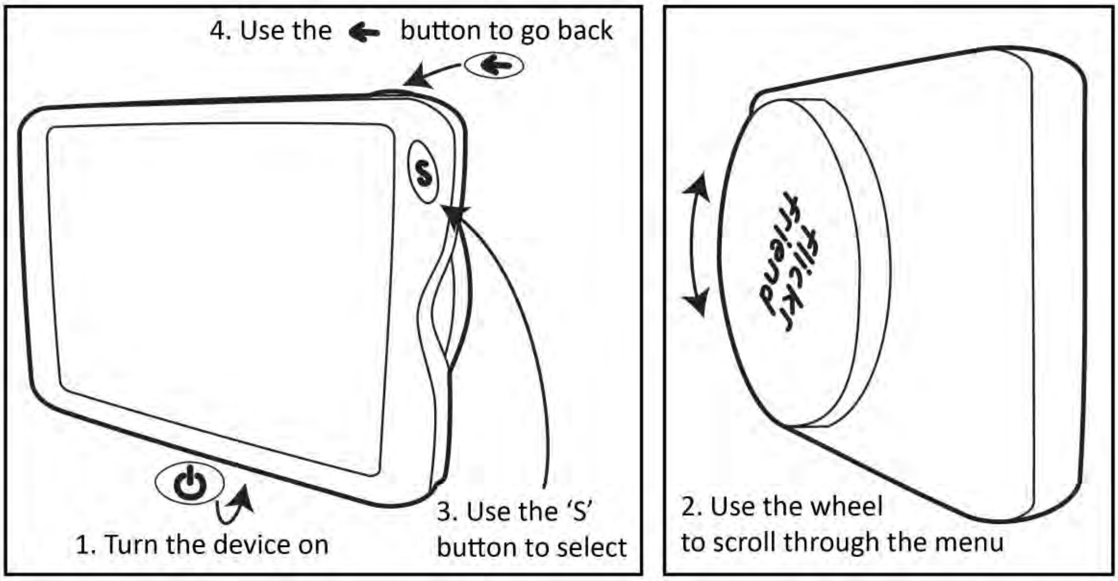


Figure 19: Interactions of the conceptual photo management product

4.4 Creation of the prototypes

Initial design work was undertaken in order to reach a stage where, in a real design process, feedback from potential users would be the next natural step. A physical form had been designed with a basic information architecture in place.

Each of the three resulting prototypes used this initial design work as the starting point. The study was intended to reflect practises of commercial projects. In these commercial projects, where ’time is money’, the level of fidelity is driven by time constraints because the time costs are the most significant cost in creating a low-fidelity prototype. Therefore time constraints were used to determine the level of fidelity of the three prototypes. Only the time to construct the prototype differed, the time constraints of 4 hours, 14 hours and 5 days were imposed. These time constraints were based on what was thought realistic based on previous experience in constructing interactive prototypes, and knowledge of the amount of time a design team might allocate to prototyping.

The designer had to decide on the best way to prototype the technical aspects within the allocated time. The considerations that have driven the fidelity level and its effects on the physicality of the model have been entirely based on these time constraints.

4.4.1 ‘Lowest Level’ prototype

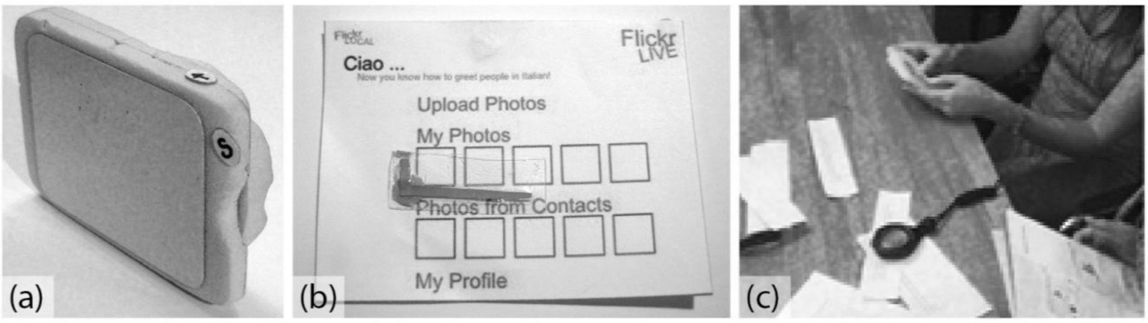


Figure 20: Low fidelity paper prototype

(a) The physical model (b) the interface (c) the trial set up

Time allowed = 4 hours (actual time taken = 3 hours 30 minutes).

Paper prototyping formed the basis of this prototype; it is a very simple technique which provides a very fast method for creating low-fidelity prototypes. For this study the paper prototype interaction was augmented by a physical prototype.

A foam model was constructed to create the physical form to scale. The foam was sanded to produce a smooth finish with white cardboard depicting the buttons and screen (Figure 20(a)).

For the digital aspect, a series of paper screens were created with a small red box to indicate which menu item is active (Figure 20(b)).

The participant held the physical model; the facilitator changed the screens and adjusted the ‘select box’ during user trials (Figure 20(c)).

4.4.2 'Mid-Level' prototype

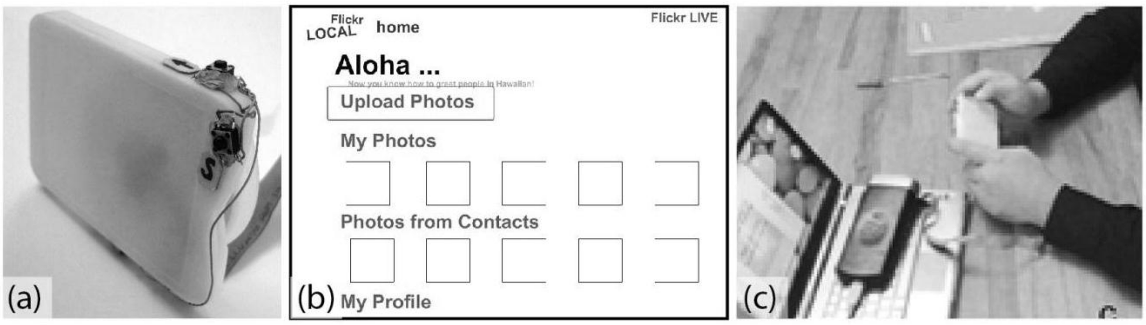


Figure 21: Medium fidelity prototype

(a) The physical model (b) the interface (c) the trial set up

Time allowed = 14 hours (actual time taken = 12 hours).

The IE System was chosen to create the mid-level prototype due to the simplicity this approach offered. The system allows a PC to receive keyboard inputs so that when a user interacts with a switch in the physical model, the PC will respond to the perceived keyboard input and a keyboard-triggered GUI is activated on the PC.

A model was created in a Computer Aided Design (CAD) system, and was constructed to scale using a Fused Deposition Modelling (FDM) machine (Figure 21(a)).

A basic menu structure was created in Adobe Flash (Figure 21(b)). The Flash animation used keyboard presses activated by off-the-shelf buttons for the screen changes, these were roughly attached onto the outside of the model and a mechanical rotary dial was glued inside the model for the 'wheel' interaction.

For the trial, the physical model was connected, with a cable, to a PC via the IE Unit (Figure 21(c)) and the visual feedback was on a laptop monitor.

4.4.3 'Highest Level' prototype

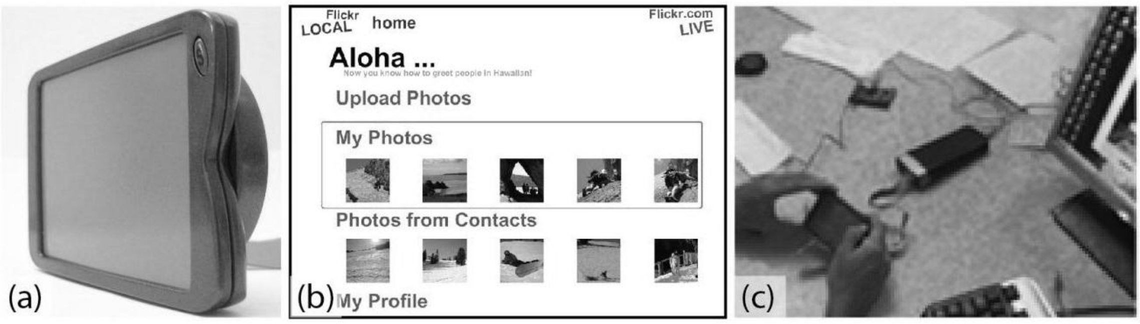


Figure 22: Highest fidelity prototype
(a) The physical model (b) the interface (c) the trial set up

Time allowed = 5 days (actual time taken = 5 days).

The extra time allowed for the highest level prototype was used to develop the following three areas: the prototype was given a realistic finish, the wheel interaction was made to feel smooth and the Flash animation was developed to operate more like the intended design. Again, a CAD model was created with design details such as shaped buttons and ports. Once the FDM model had been made, it was sanded and sprayed (Figure 22(a)). Dome switches that produce positive tactile feedback with a low profile were used for the buttons triggering the Flash animation through the IE Unit. The analogue dial was an off-the-shelf Phidget component. This reflected the intended physical-digital interaction of the design intent better than the rotary dial used in the mid-level prototype. The Flash animation had more realistic menus and a smoother transition between screens (Figure 22(b)). For this trial, the physical model needed to be connected through both an IE Unit and a Phidget Interface Kit with wires, and the visual feedback was on a desktop monitor (Figure 22(c)).

4.5 Comparative Analysis of the Prototypes

The resulting prototypes differed considerably, and their properties are reviewed in relation to McCurdy *et al.*'s (2006) five dimensions of fidelity, as shown in Table 2. A similar technique is applied in Table 3 to analyse the subsequent effects on physicality, which are considered to fall under two areas: the physicality of the product itself (e.g. form, finish, weight) and the physicality of the interaction (feel of the buttons and wheel in this case).

Table 2: Properties of each prototype in relation to the Five Dimensions of Fidelity (McCurdy, Connors, Pyrzak, Kanefsky, & Vera, 2006).

Dimension of fidelity	Driving factors	Lowest level 3 hours 30min	Mid-level 12 hours	Highest level 5 days
Aesthetics	Model material	**Blue foam** (both material and finish differ considerably from intended design)	**Unfinished FDM** (similar material but finish differs considerably from intended design)	**Sanded & sprayed FDM** (similar material and finish to intended design)
	Model finish			
Richness of interactivity	Wheel mechanism	**Free rotating** (similar to intended design but no real-time feedback given)	**'Clunky', clicking mechanism with end points** (very different from intended design but gives real-time feedback)	**Smooth mechanism with end points** (very similar to intended design and gives real-time feedback)
	Buttons	**Cardboard representations** (very different in feel and aesthetics from intended design)	**Switches tacked onto model** (very different to intended design but gives real-time feedback)	**Integrated switches** (very similar to intended design in look and feel gives real-time feedback)
	Screen operation	**Paper screens** (no real-time feedback so very different from intended design)	**Basic Flash animation** (real-time feedback but sketchy interface, differs slightly from intended design)	**More advanced Flash animation** (real-time feedback and graphics similar to intended design)
Depth of functionality	Screen operation	All have identical features enabled, feature will appear 'unavailable' if it is not part of a task		
Breadth of Functionality	Screen operation	All have identical menu structures, the tasks chosen highlighted the breadth of functionality in the intended design		
Richness of Data	Data used	**Sketch data used** (different from intended design)	**Sketch data used** (different from intended design)	**Photos used** (very similar to intended design)

Table 3: Properties of each prototype in relation to the areas of physicality

Area of Physicality	Driving factors	Lowest level 3hrs 30min	Mid-level 12 hours	Highest level 5 days
Physicality of the product	Scale Model material Screen material Weight	**1:1, made from blue foam with a cardboard screen** (form is very similar to intended design, finish and weight is considerably different)	**1:1, unfinished FDM with screen placement suggested on model (no colour difference)** (form is very similar to intended design, weight and finish are considerably different)	**1:1, finished and sprayed FDM with a colour difference depicting the screen** (form and surface finish is very similar to intended design, weight is different)
Physicality of the interaction	Wheel mechanism	**Wheel freely rotates** (as intended in design) **with no real-time physical or digital feedback** (extremely different from intended design)	**Mechanism feels clunky and cannot rotate continuously** (considerably different from intended design) **gives real-time physical** (not part of intended design) **and digital feedback** (part of intended design)	**Mechanism feels smooth** (very similar to intended design), **cannot rotate continuously** (not part of intended design) **gives real-time physical and digital feedback** (similar to intended design)
	Buttons	**Buttons are depicted with cardboard and give no physical or digital feedback** (very different to intended design)	**Buttons are off-the-shelf and tacked onto the model** (very different to intended design) **but give real-time physical and digital feedback** (similar to intended design)	**Buttons are integrated dome switches with real-time digital and physical feedback** (very similar to intended design)

4.6 The User Study

4.6.1 Location

All trials were conducted within an 'ad hoc' lab; this included a video camera on a tripod which captured visual and audio data.

4.6.2 Participants

A pilot study was first carried out with 9 undergraduate participants from the university.

The main study was conducted using 48 participants recruited from UWIC (now Cardiff Metropolitan University) staff who have used digital cameras (including cameras on their mobile phones). The participants were randomly divided into three independent groups of 16, one for each fidelity level, to eliminate possible learning effects. There were 23 females and 25 males, with ages ranging from 19 to 50, with an average age of 29.

In order to ensure all participants were 'potential users', a participant screening document is needed. The product under trial is a main-stream consumer electronics product designed for a wide variety of users, therefore screening was kept to a minimum. Participants were excluded if they had any previous knowledge of the research.

4.6.3 Structure

The following structure was applied to every participant for each of the three prototypes trialled:

1. Participant fills in a demographic questionnaire covering age and gender, and indicates existing technology usage. (Note the prototype is not in sight at this stage.)
2. Participant is given a written description of the product.
3. Facilitator uncovers the model and records if the participant picks it up, and the reaction in relation to the fidelity of the aesthetics.
4. Participant is given five tasks to carry out. Moderator records whether the user experienced a success, minor problem, serious problem, or a catastrophe.
5. Participant fills in a questionnaire, and is asked to rate certain aspects of their experience with the product.

4.6.4 Participant Information Sheet

The participant is given some information about the study to read through before signing a consent form. The information sheet can be found in Appendix 7.

4.6.5 Pre-task questionnaire

The questionnaire is shown in Appendix 7. The moderator completed the user number then handed the sheet to the participant to complete the remainder.

4.6.6 Written description

Once the questionnaire had been completed, the participant was given a written description of the product:

> *"The model that you are about to be shown is a prototype of the Flickr Friend. The concept is based around the Flickr website. It is intended as a mobile wireless viewer of your online Flickr pages, additionally it has a hard drive so that you can store and manage all your photos and choose if they will be viewable online. The Flickr Friend also has a USB slot for attaching a digital camera."*

After reading the description, the moderator asked a series of questions to ensure the participant fully understood the concept of the product and the Flickr website.

4.6.7 Tasks

The set of trials and rating scale used to classify the severity of problems was based on recommendations by Redish *et al*. (2002). Each participant was given a series of 5 scripted tasks:

1. Turn the product on
2. Find a photo on the *Flickr* website:

"You would like to show a friend a photo that you have recently uploaded to your Flickr webpage. Locate the photo of a climber"

3. Find a friends photo on the *Flickr* website:

"Your contact 'Red Rocket' has just uploaded photos from your last trip (snowboarding). Find the photos and view them as a slideshow."

4. Find a photo from the hard-drive:

"You remember a photo of Guy Fawkes located on your local drive that you took on Bonfire Night last year. Upload the photo to Flickr for others to see."

5. Transfer a photo from a camera:

"You are on holiday and would like a photo from a friend's camera. Locate the photo of a beach and transfer it to your hard-drive."

4.6.8 Post-task questionnaire

The participants were then asked to complete an 8 point scale ('best' to 'worst') in the following areas:

- The look of the product
- The feel of the product
- The feel of the interactions
- Their initial emotional response
- Ease of use

4.7 Data Captured

A mixture of quantitative and qualitative data was gathered from the study for analysis. This section describes how the quantitative and qualitative data was captured.

4.7.1 Quantitative Data

The 'performance rating' was captured during the study. The usability problems that caused users to struggle with individual tasks were also recorded for analysis. All ratings and usability problems were subsequently reviewed using video capture data to ensure consistency with the coding scheme.

4.7.1.1 *Performance rating*

The moderator was responsible for rating a user's performance for each task on the basis of the following criteria. These categories are based on those proposed by Molich & Dumas (2008) for reporting usability problems:

A Catastrophe:

- The user is unable to complete the task.

- The user refuses to complete the task.
- The user solves the task incorrectly without noticing.

For example: the user has not completed a task (even if he/she thinks they have) or the user gives up.

Serious problem:

- The user is significantly delayed (1 – 5 minutes) but manages to complete the task.

For example: the user repeatedly tries the incorrect menus or buttons.

Minor problem:

- The user is briefly delayed; the user experiences a problem, but corrects themselves reasonably quickly (less than 1 minute).

For example: the user goes into the wrong menu, user cannot find a button.

Success:

- The user completed the task without problems or delay.

For example: the user finds all the correct buttons and menus when needed.

4.7.2 Qualitative Data

Qualitative data was also gathered through a semi-structured interview at the end of the test and comments and user interactions were captured on video during the test.

A digital video (DV) camera was used to capture each of the participants' sessions; this was transcribed and then coded thematically. Video clips were managed using the qualitative video analysis software *Transana* (Woods & Fassnacht, 2007).

4.8 Comparative Analysis of the Prototypes in the User Study

This section describes how both the quantitative and qualitative data was analysed.

4.8.1 Quantitative Analysis

The 'performance rating' data shows whether the task was a success, had minor or major problems or was a catastrophe. The performance data was converted into interval data (3 = success; 2 = minor problem; 1 = major problem; 0 = catastrophe) and analysis was conducted using a 3 (prototype level) by 5 (tasks) mixed analysis of variance (ANOVA) with the alpha level set to 0.05.

The ANOVA results show a non-significant effect of the prototype on the task outcome $F(1,45) = 1.66$, $p = .201$. Figure 23 shows the results graphically, including error bars.

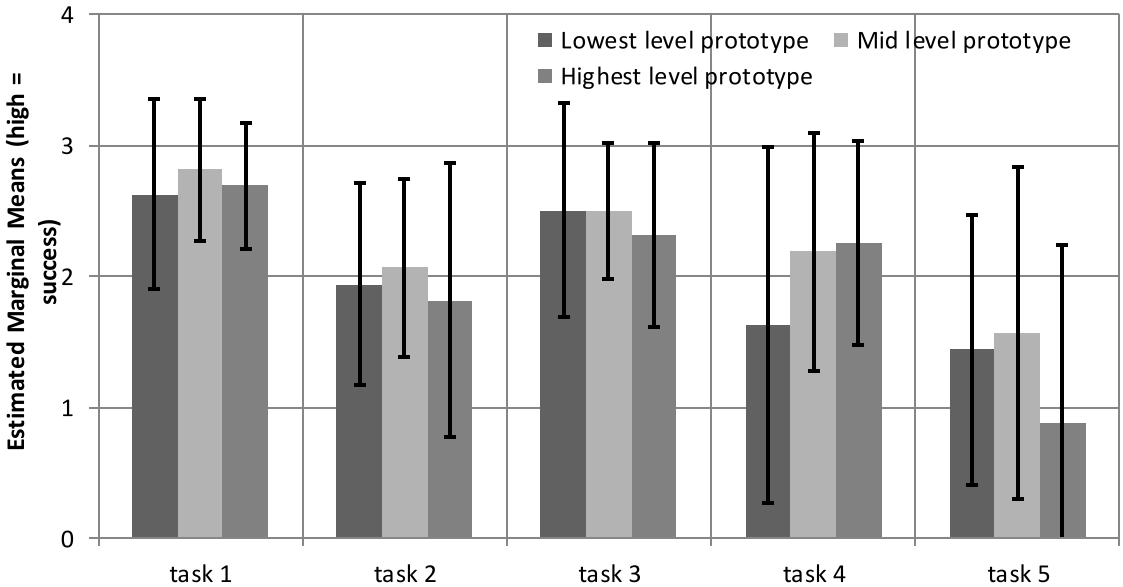


Figure 23: Performance ratings for each of the 5 tasks as a function of product type

The quantitative data on its own did not reveal any differences, which indicates that neither differences in physicality nor in fidelity have an effect in this particular study, or that this is not a reliable way of analysing this effect. The qualitative data was subsequently analysed to determine if further effects of fidelity and physicality could be uncovered.

4.8.2 Qualitative Analysis

The qualitative analysis was conducted by reviewing the video recordings of each participant after the trials. The qualitative analysis was in two stages: firstly, identifying problems that participants may have encountered while performing each task (Part 1

Analysis) and secondly, assessing whether participants were influenced by the fidelity and physicality of the prototypes (Part 2 Analysis).

Part 1 Analysis: This was conducted to find out where participants were having problems when performing each task (*types of usability problems)*.

During the trials, the main errors were observed and noted in a table. When reviewing the video data, each error made by the participant was recorded and any additional errors were included. The errors were then condensed into three problem areas, which were identified as being of hindrance to a user in completing a task. The problems areas were:

a. Unclear meanings of symbols (semantics)
   For example; user presses a button but comments that they do not know where it will take them because of the button symbol.
b. Difficulty locating appropriate interface elements (interactions)
   For example; user comments that they need to scroll through a menu but cannot find the scroll wheel.
c. Unexpected feedback from software (mental model mismatch with the information architecture)
   For example; user expresses a surprise when they navigate somewhere in error.

If a participant kept repeating the same error, it was recorded several times, this highlighted particular areas of concern. The results of the analysis are shown in Figure 24.

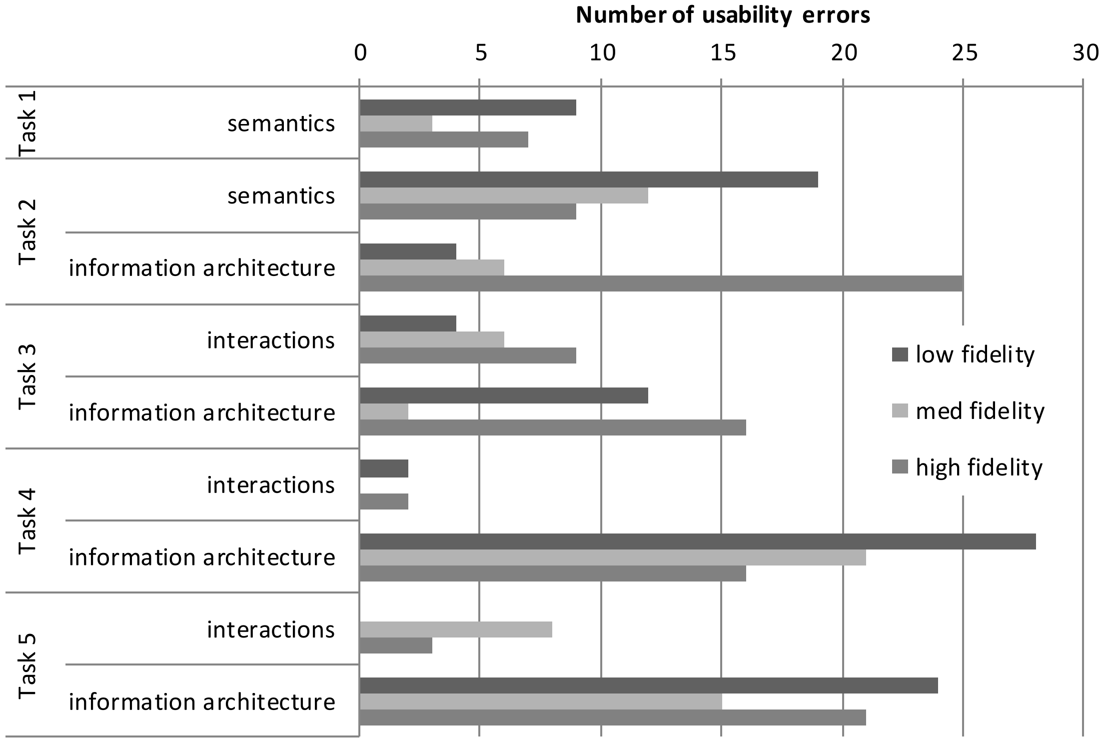


Figure 24: Usability errors coded by the three areas of investigation

The results that are of particular interest are those that differ across the prototypes. So, for example, during Task 2 there were 4 problems recorded by the lowest level prototype due to the 'information architecture' but the same problem area resulted in 25 problems for the highest level prototype. Other notable results are again for Task 2, where users had problems with the 'semantics' 19 times for the lowest level, 12 times for the mid-level and 9 times for the highest level. For Task 5, there were 0 problems for 'interactions' for the lowest level prototype, but 8 problems for the mid-level and 3 problems for the highest level.

Part 2 Analysis: This was undertaken to assess whether participants were affected by the fidelity and physicality of the prototypes based on the comments made, for example, '*the wheel mapping is not natural*'.

A similar recording procedure was followed as in Part 1 Analysis using the errors noted during the trials plus the video review. The comments were then coded and those relating to the following themes were selected and collated into groups:

1. physicality of the product (e.g. size in the hand, screen position and comfort)
2. physicality of the interaction (e.g. the button is in the wrong place, how the wheel feels)

3. feedback about the design and idea in general.

For the example given above, which concerns the wheel mapping, the comment would be coded under 'physicality of the interaction'. The results are shown in Figure 25.

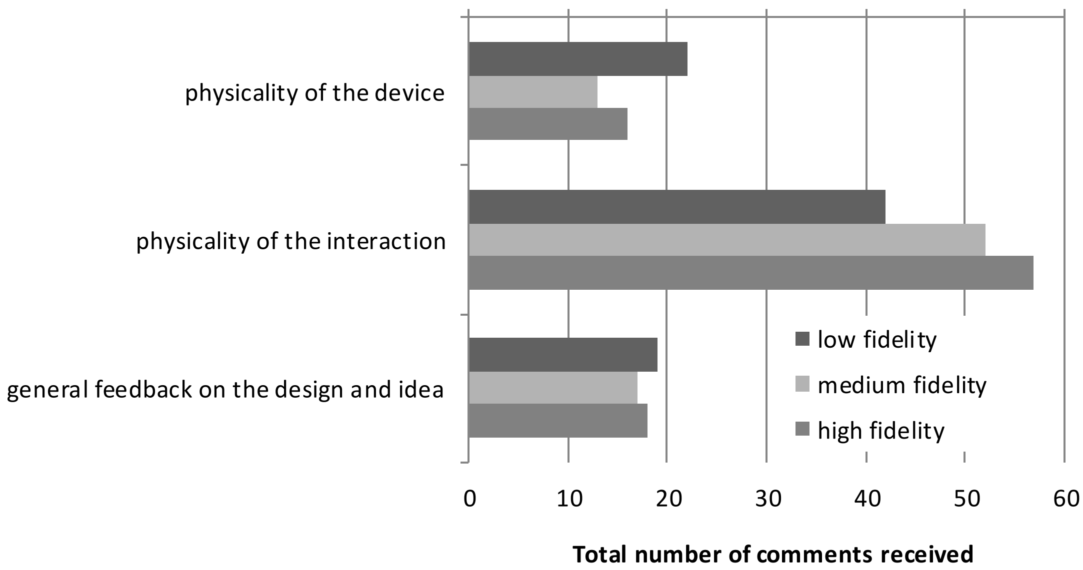


Figure 25: The number of comments received for each area of investigation

The general feedback on the design and idea is roughly the same across the prototypes. The lowest level prototype seems to differ from the mid-level and highest level prototypes. The physicality of the *product* received 22 comments for the lowest level prototype, compared to 13 at the mid-level, and 16 at the highest level. The physicality of the *interaction* received 42 comments for the lowest level prototype, compared to 52 at the mid-level, and 57 at the highest level.

4.9 Discussion

The analysis produced a number of interesting results that require further discussion.

The results of the quantitative analysis show that there is little difference in performance across the prototypes with different fidelity levels (which would seem to agree with the research by Lim *et al*. (2006)). This, in itself, is an important result showing that in the early stages of the design process, the fidelity level might not have a significant impact. Despite the mid-level prototype being physically different from the intended design in a number of seemingly important ways (the wheel clicked, could not rotate 360° and felt very 'clunky'), it still produced valid feedback about the concept. Furthermore, the mid-level prototype took less than half the time to build compared to the highest level prototype. Even the lowest level paper prototype seemed to produce usability data in line with the higher fidelity ones.

As mentioned in Part 1 of the qualitative analysis, and shown in Figure 24, the results for the problem area 'information architecture' in Task 2 differ across prototypes, with 4 problems recorded for the lowest level, 6 for the mid-level and 25 for the highest level prototype. When combined with the rest of the data from the analysis, suggestions can be made about the cause of this difference. One explanation could be that users of the lowest level prototype had so much trouble with 'semantics' for both Tasks 1 and 2 (28 problems in total) that there were very few problems with the 'information architecture' (4 problems) for Task 2. Compare this to the highest level prototype, where users had fewer problems with 'semantics' in Tasks 1 and 2 (18 in total), but they had difficulty with the 'information architecture' of the product (25 problems). The mid-level prototype has the lowest number of problems related to 'semantics' (12 in total), and 6 problems concerning the 'information architecture'. The difference in results for 'semantics' is an interesting one because this is down to interpretation of the symbols, however, the symbols were identical across the three prototypes. These results seem to indicate a correlation between the 'semantics' and the 'information architecture'. This may be because many users worked out what interactions did by 'experimenting' with them instead of understanding the symbols, and thus creating their own mental model based on the feedback they received on screen. This explorative approach was only supported by those prototypes that had real time feedback, the lowest level paper prototype required the facilitator to find the correct feedback and change the screen, making the feedback one step removed from the interaction. The indication from this is that users of the mid and highest level prototypes, with real time tactile and digital feedback (or 'automatic' feedback), had fewer problems with 'semantics' but subsequently more problems with the 'information architecture'.

Later on in the tasks, the situation seems to change. The results of Task 4 show that users of the lowest fidelity prototype had the most problems with the 'information architecture' (28), whereas there were 21 problems for the mid-fidelity prototype and 16 for the highest fidelity prototype. The large number of problems for the lowest fidelity prototypes could be because users of that prototype were so distracted by their problems with the 'semantics' early in the tasks, that they had not formed good mental models of the information architecture. This could possibly be due to the users' inability to fully engage with the product and therefore following a 'more luck than judgment' approach to interaction. This indicates that users of the mid and highest level prototypes had more problems with the mental model of the product early on in the trial whereas the lowest level prototype users encountered these issues later on in the trial.

However, there is one anomaly in the results presented. It has been proposed that the automatic prototypes had fewer problems with 'semantics' but subsequently more problems with the 'information architecture'. However, the mid-level prototype only elicited 6 problems due to the information architecture (compared to 25 for the highest-level prototype and 4 for the lowest), which despite being more than the low fidelity prototype, is still relatively low. Perhaps this could be down to the feel of the interactions themselves. When the physicality of the prototype is taken into account (Table 2 on page 87), it can be seen that the lowest level prototype has no tactile feedback on pressing the buttons (just the facilitator moving a screen), while the mid-level prototype has very pronounced buttons that give both tactile and on screen feedback, and the highest level prototype has more subtle visual properties with subtle tactile feedback plus on screen feedback. The indication from this is that the more 'positive' feel of the buttons aids the user in navigating an interface on the prototype. Indeed, this seems to hold true for all the tasks (except Task 5 which will be discussed next). In all tasks participants using the mid-level prototype had less problems in all areas than the other two prototype.

The results for Task 5 do not seem to match up with any of the theories proposed so far. Task 5 involved the connection of an external product (a camera) to the prototype. The first four tasks were predominantly based on the interface and button-based interactions, instead, the fifth task relied on the participants noticing how to physically connect a camera. On reflection, this task did not allow participants to use their experience of the first four tasks in order to help them complete this final task. Therefore the results of Task 5 are not necessarily indicative of the participants' development with the product.

The results from part 2 of the qualitative analysis show that more comments were received about the physicality of the interaction than the physicality of the product. This suggests that the test was set up in a way that elicits more comments about the physicality of the interaction rather than the physicality of the product. However, the lowest level prototype received more comments about the physicality of the product unlike the mid and highest levels and fewer comments about the physicality of the interaction. This could be because the physicality of the interaction was so far removed in the lowest level prototype from that which was intended, hence it was harder for users to judge this aspect of the design and, as a result, they made more comments about the physicality of the product itself.

Finally, the comparative analysis of fidelity and physicality of the prototypes was not able to determine a correlation between fidelity and physicality, therefore it cannot be determined that higher fidelity prototypes result in higher levels of physicality, for example. This could be because of the way in which physicality was assessed; the approach used provided a text-based overview of each prototype, but it does not give an indication of the 'level' of physicality.

4.10 Conclusion

This study has reported on an exploratory investigation into the effects of physicality and fidelity on the prototypes used for front-end product design development. Each of the prototypes created represented the same design intent, and enabled the same functionality. Time constraints governed the fidelity level, and each prototype was tested for its 'level' of physicality and number of usability issues.

The trials suggest that there is no effect of fidelity and physicality at the early stage of the design process in terms of user performance; however, combined with the quantitative analysis points of interest can be seen:

1. All prototypes achieved similar results for the performance test.
2. Users of the mid and highest level prototypes, with real time tactile and digital (on screen) feedback, had fewer problems with semantics.
3. Users of the mid and highest level prototypes had more problems with the information architecture of the product early on in the trial whereas the lowest level prototype users encountered these issues later on in the trial.
4. The mid and highest level prototypes gave more feedback about the physicality of the interaction.

These results suggest that for the initial exploration of a design idea; very low fidelity prototyping is a fast and low cost method of getting reliable feedback especially in relation to information architecture. Alternatively, if more specific feedback about the intended design and interaction is required, then a prototype that can produce immediate feedback is essential. However, there are many more factors at play, and these need to be explored further to inform design guidelines in relation to the needs of the early design process.

The initial hypothesis of this study was that '*a relationship exists between fidelity and physicality and this relationship affects the effectiveness of the prototype*'. Some of the results do indicate a correlation between fidelity, physicality and the outcome of the user trial; however, some of the results indicate there is no correlation. This leads to the conclusion that the hypothesis was perhaps too generalised and attempted to over-simplify the situation. The assumed link between fidelity and physicality could be the cause of this. The hypothesis was based on the presumption that the level of fidelity would affect the level of physicality, and time constraints placed on the creation of the prototypes controlled the level of fidelity. However, the comparative analysis of physicality was not able to reveal any insights or correlation between fidelity and physicality. This indicates a need for a common framework by which to assess and compare physicality, and this is developed further in the next section.

4.11 Development of a New Construct for Physicality

Study One sought to uncover the resulting differences in physicality based on low, medium and high(er) fidelity prototypes. However, the approach taken for a comparative analysis was not able to uncover a correlation between fidelity and physicality. This was proposed to be because the 'level' of physicality could not be determined in the descriptive approach used. In Study One, physicality was considered to fall under two areas: the physicality of the product (for example; form, finish and weight) and the physicality of the interaction (the feel of the buttons and wheel in this case). But this approach only allows the prototypes to be described and compared to one another in such a way that would indicate that one has 'more physicality than another'. A comparative approach is essential when using physicality to determine the differences between the prototypes on trial.

The two 'areas' of physicality (the product and interaction), used in Study One, were used as a starting point for development. Initially, an attempt was made to 'score' the prototypes constructed for Study One in order to compare them. However, this proved difficult because the two 'areas' of physicality lack a clear distinction between the 'product' and the 'interaction'. However, concurrent work on 'Physigrams' (presented in the Contextual Studies on page **Error! Bookmark not defined.**) uses the notion of the product 'unplugged' in order to separate the physical interaction elements of the design from the electronic effects of those interactions. This way of interpreting the prototype provided a valuable starting point to separate the physicality of the 'product' from the physicality of the 'interaction'.

Central to the Physigram study, was the ability to consider the product 'unplugged' and entirely separate from any digital or other external functionality. This was used in order to focus solely on formally capturing the physical aspects of an interaction. However, if an interactive prototype is considered 'unplugged' in the same manner; the physicality of the 'product' can be determined to be related to the physical properties of the object, whereas the physicality of the 'interactions' are the physical properties of the experience of interacting with the prototype. This consideration of the prototype 'unplugged' lead to the idea that physicality in relation to the product is 'passive' because it relates to the static object, and that the physicality of the interaction is 'active' because it requires active engagement with the prototype.

Thus the constructions of 'passive' and 'active' physicality were developed where; passive physicality is proposed to be the **perceived affordance based on the tangibility and aesthetics of the prototype,** and active physicality is proposed to be the **tangible and visual experience of interacting with the prototype.**

To explain the notion of passive physicality further, consider the product 'unplugged'; passive physicality is concerned with the judgments that can be made about a product by considering both its tangibility (by touching it), and its aesthetics, without switching it on. Assumptions are formed about the physicality of the product based purely on its aesthetics, as Reeves (2006) demonstrates by asking; do you grasp a cup by its handle or by the body? Decisions are made about the comfort of the cup's handle by its appearance and the perceived weight of the contents of the cup. Passive physicality also has its roots in Gibson's description of affordances (1977), and discussed in detail in Section 2.2.2 of the literature review. Affordances suggest ways of interacting with an object. They are not simply a property of the object; they are the way that a specific user relates to that object. When Norman (1998) applied Gibson's idea to design, he divided the idea of affordances into those of real and perceived affordances. Whilst real affordances are what the user could actually do with a product, meaningful or not, perceived affordances tell the user '*what actions can be performed on an object and, to some extent, how to do them'*. The design of the product has affordances; in addition, the way in which the prototype is constructed brings its own, different, affordances that affect the way in which the user perceives the object. Passive physicality forces the designer to recognise that the way in which the physical prototype is executed has a significant impact on the user's experience of that prototype.

Active physicality, on the other hand, is concerned with the physical act of interacting with a prototype in its 'on' state. This interaction results in both tangible and visual feedback; for computer embedded products the tangibility of the interaction would be meaningless without the feedback of the interface operating in a realistic manner, be it a graphical interface, light or a mechanism. Therefore, active physicality is the combination of both the tangible and visual feedback of the interface.

This description of active and passive physicality brings together the three philosophical discussion areas related to physicality and the designed object as identified by the literature review. These three areas were: humans as physical beings within our physical world (embodiment), physical signifiers for interaction (affordances) and the point at which the digital and physical meet (interaction). The proposed definition of active and passive physicality recognises that, as individuals, we bring our own understanding and experiences with us when we interact with a product; we are embodied within our physical world. This influences how affordances of the prototype are interpreted, an affordance can exist but without knowledge of how to act on that affordance it might not be perceived correctly. This is especially true of prototypes where the way in which the product is realised will differ, sometimes considerably, from the intended design. And finally, by separating active physicality and passive physicality, this notion addresses the very point at which the interaction between the digital and physical elements of the design come together. The notion of active and passive physicality intends to give designers (and those constructing interactive prototypes) confidence that any prototype created will produce meaningful data in a user trial.

The second study presented in this thesis is intended to explore the notion of passive and active physicality, as described here, in more depth. In addition, the two user studies that have already been presented in this thesis will be re-examined to see if these notions hold true and perhaps offer clarity to the findings of the contextual study and Study One.

Chapter 5. Study Two: Media Player

5.1 Introduction

Study One, presented in Chapter 4, sought to uncover the resulting differences in physicality based on low, medium and high(er) fidelity prototypes. In Study One, physicality was considered to fall under two areas: the physicality of the product (e.g. form, finish, weight) and the physicality of the interaction (the feel of the buttons and wheel in this case). Chapter 4 discussed how this framework was developed further to form the notion of passive and active physicality where;

- Passive physicality is the **perceived affordance based on the tangibility and aesthetics of the prototype.**
- Active physicality is the **tangible and visual experience of interacting with the prototype.**

The hypothesis of Study Two was:

> *The level of physicality can be described using the framework of active and passive physicality and the level of physicality will influence the effectiveness of the prototype.*

By attempting to understand physicality, and using this to drive the physicality of low fidelity prototypes, this study aims to draw out just how physicality can be used by the designer to create efficient low fidelity prototypes.

The following paper was published and presented at Interact 2013 and can be found in Appendix 8.

> Hare, Gill, Loudon & Lewis (2013), **The effect of physicality on low fidelity interactive prototyping for design practice**. In the proceedings of Human-Computer Interaction - Interact 2013, Cape Town, South Africa. Springer Berlin Heidelberg.

5.2 Method

As discussed in the Methodology Chapter, a two-stage approach was taken to the research studies; the first stage was a comparative analysis of the physicality of the prototypes created and the second stage was a comparative analysis of the prototypes through user trials.

This two-stage approach sat within the wider study method which consisted of the following:

1) Choosing the product to trial
2) Generation of the prototypes
3) Comparative analysis of the physicality of the prototypes
4) Conduct the user trials
5) Comparative analysis of the prototypes through the user trials

The next sections provide further detail for each of these steps.

5.3 The Product under Trial

In contrast to Study One, an existing product was chosen for Study Two to provide a datum against which the retrospectively developed prototypes could be measured. The choice to retro-prototype an existing product as a method was taken after considerable thought, as discussed in Chapter 3. For Study Two, retro-prototyping was chosen because it has the benefit of access to a real, mass produced product, identified by the manufacturer as a worthwhile idea and having successfully undergone a product development process. The finished product can be used to compare the results from the user study in a manner that is all but impossible to recreate in a research study.

The decision concerning what product to prototype was based on a number of factors. The primary considerations were for a handheld product which had physical buttons and had a wide market reach (to aid in participant recruitment). An additional consideration was for it to include a non-button physical interaction such as a dial or slider in order to explore a wider range of inputs.

Figure 26: The iRiver SPINN

The product chosen was the iRiver Spinn (Figure 26), a personal media player released for sale in January 2009. It has a large screen on the front, with a dial on the side; this dial is called the 'Spinn'. There are no buttons visible from the front but there are two buttons on the top edge and a further two on the side. Figure 27 shows the interactions with the product. The user first turns the product on with the button on the side. The dial is used to scroll between menu items and the selected menu item can be selected by pressing down the Spinn dial. Volume is controlled from the side and the 'back' button is located on the top edge of the product.

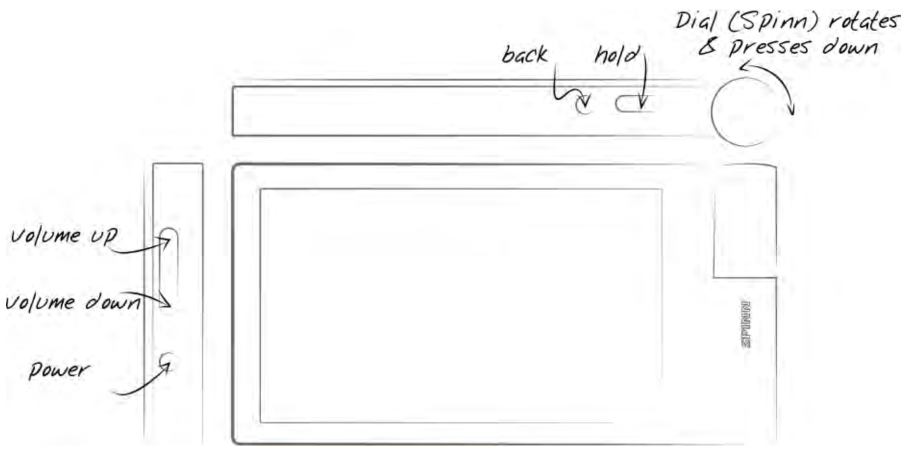


Figure 27: The interactions of the iRiver Spinn

Engadget reviewed the iRiver Spinn in 2009 (Miller, 2009) concluding with:

"It's plain to see that iRiver has obviously put a lot of thought into the design of this player, we just wish they'd put half as much thought into everyday usability. There's so much to love about the SPINN, and a bit of an x-factor that makes us want to love it more, but when we really sit down and try to build our portable media life around this little flash-based (player), there are too many drawbacks in actual usage to make it worth it."

This trial is intended to identify some of those ‘drawbacks’ or usability issues through a series of low-fidelity prototypes. The article by Engadget is used later as an ‘expert review’ by which the results of the user study could be validated.

The product is intended to be touch screen, but this capability was not implemented in the prototypes. This is because the study design sought to focus on the physical interactions with the product and the interaction was designed to be possible without using the touch screen. Upon further investigation, the screen was identified as not being very responsive (in a video review (Digital Trends, 2009), even the promoter has to touch the on-screen ‘back’ button five times for it to work). Engadet say of the touchscreen: “*The product doesn't have any touch-and-swipe motions, you'll have to grab the scroll bar and pull, so it usually makes sense to just spin. When you do pull and drag, or try and tap items, the product somehow seems to act slower than it does when you work with the spin wheel*”. All participants were requested to use the prototypes through the physical inputs (in fact, in the user trials some participants of the final product did try and touch the screen, but a lack of response pushed them towards the physical controls naturally).

5.4 Creation of the Prototypes

Four low fidelity prototypes were constructed using techniques identified in the literature review. Each prototype was planned giving due consideration to active and passive physicality levels, with the intention of placing one in each of the quadrants shown in Figure 28 below.

A single interface was coded in Adobe Flash for all prototypes and adapted to the needs of each. Preparatory work ensured that this interface would be suitable for all prototypes, and that the adaptation of the interface was possible for all. As is typical at this stage of the design process, only a limited selection of features were included in the software (Nielsen J. , 1993). A single Computer Aided Design (CAD) model was created.

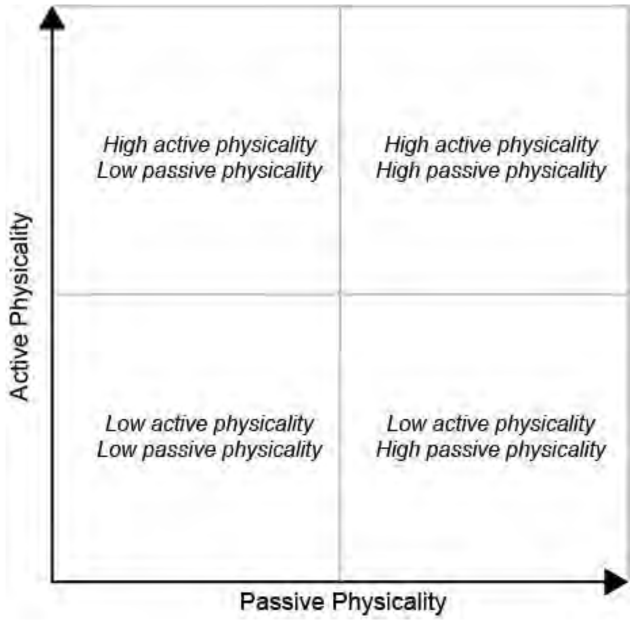


Figure 28: Areas of physicality

5.4.1 The 'blue foam' prototype

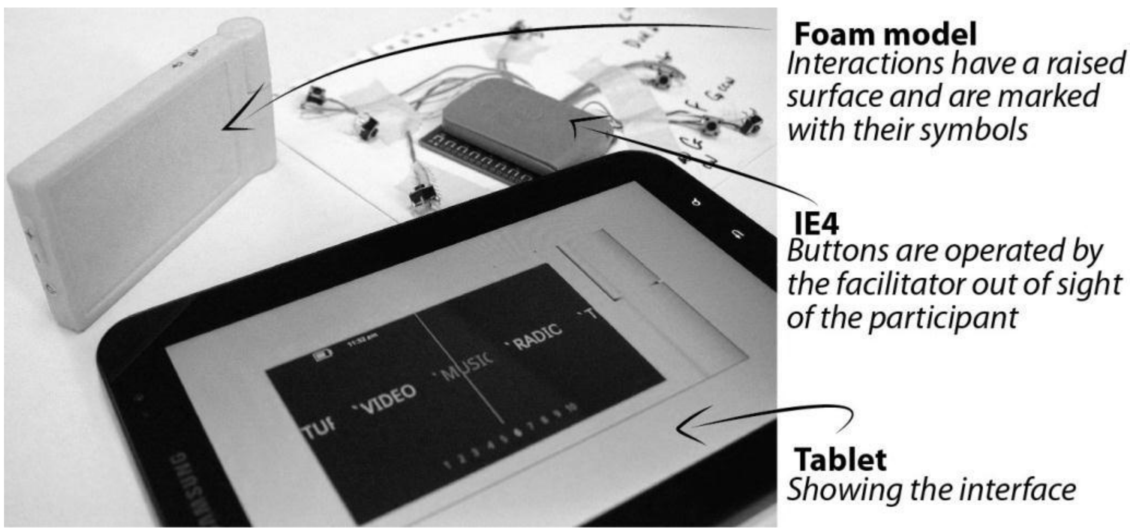


Figure 29: The blue foam prototype

The **'blue foam'** prototype (Figure 29) was constructed from model making foam board; the model had no electronics or buttons embedded within it. The interface was shown on an Android tablet displaying a Flash representation of the interface. The facilitator operated the interface wirelessly with an IE4 configured to send keyboard triggers to the tablet through Bluetooth. The participant was asked to interact with the foam model and follow the 'think out loud' protocol (Gould & Lewis, 1983), so that the facilitator could operate the interface based on what the participant was saying and interacting with on the foam prototype.

5.4.2 The 'white model' prototype

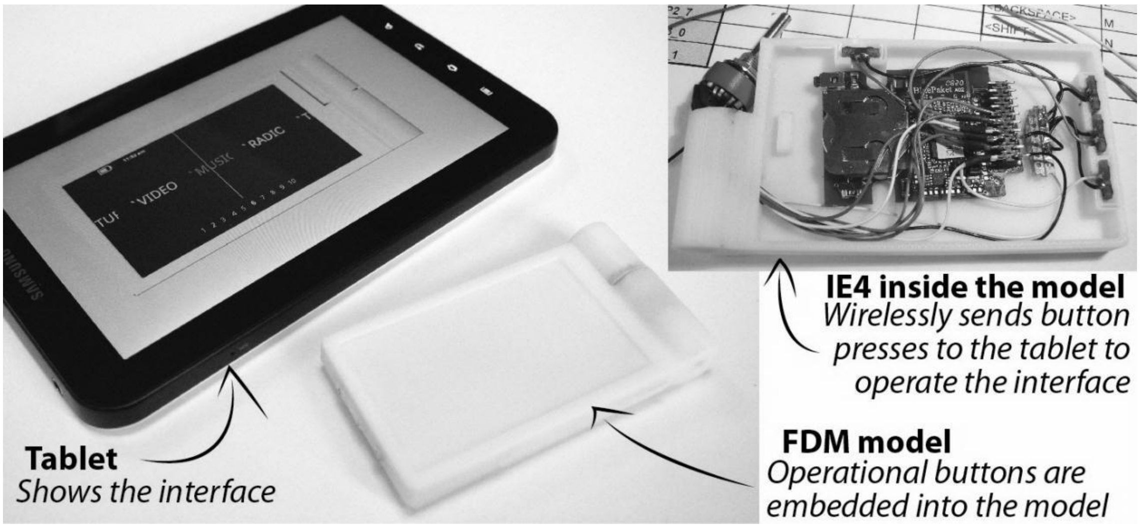


Figure 30: The white model prototype

The physical model for the **'white model'** (Figure 30) was created in CAD and constructed using rapid prototyping techniques (FDM). The CAD model was adjusted slightly to house the buttons and the dial which were integrated to make the prototype interactive. An IE4 (Gill, 2013) was used to connect the buttons to a laptop. The Flash interface, shown on a tablet, 'listens' for key presses from the IE4, and triggers changes in the interface when the participant interacts with the prototype. The dial component had 12 segments which were wired to seperate key presses and coded to represent the dial turning.

5.4.3 The 'appearance model' prototype

The physical model for the '**appearance model**' (Figure 31) was intended to reflect the final product as accurately as possible. The physical model was created on a rapid prototyping machine, using FDM identical to the 'white model', and then finished to facsimile level. Buttons were integrated into this model but in order to keep the size identical to that of the final product, the buttons did not function. For this prototype, participants' could use the physical model to get a feel of the design but all interaction was performed on a tablet through its touch screen. The dial interaction was simulated by buttons representing clockwise and anti-clockwise turns, each press of a button scrolled the menu on by one item.

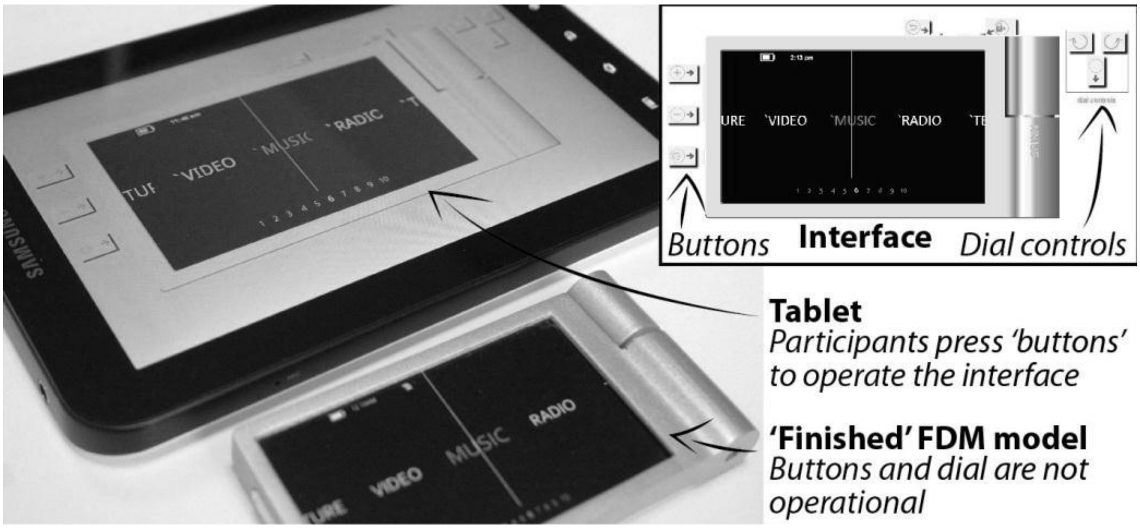


Figure 31: The appearance model

5.4.4 The 'blue foam with wires' prototype

A rough foam model was constructed for the **'blue foam with wires'** (Figure 32) to accommodate the off-the-shelf buttons and dial. The dial was connected to an Arduino (Burleson, Jensen, Raaschou, & Frohold, 2007) which received the analogue signals and outputted them to the computer running the Flash interface. The buttons were connected to an IE4. As a result of the extra code required for the Arduino, the interface was shown on a laptop rather than the touch screen tablet.

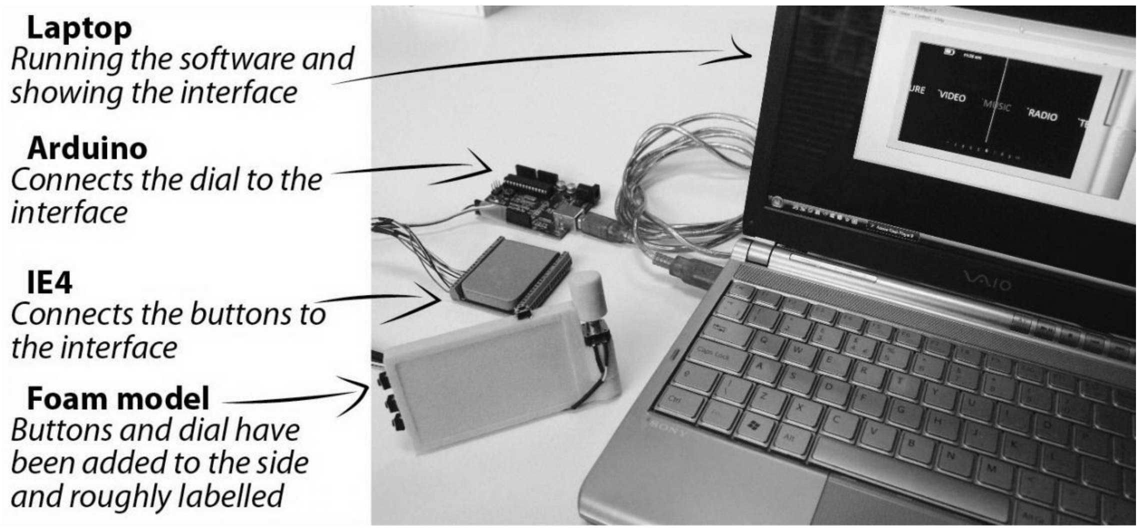


Figure 32: The blue foam with wires

5.4.5 Time taken (and the subsequent cost of construction)

The time taken to construct each prototype was recorded during the construction phase; this was divided into the physical model and the interface, this is shown in Figure 33.

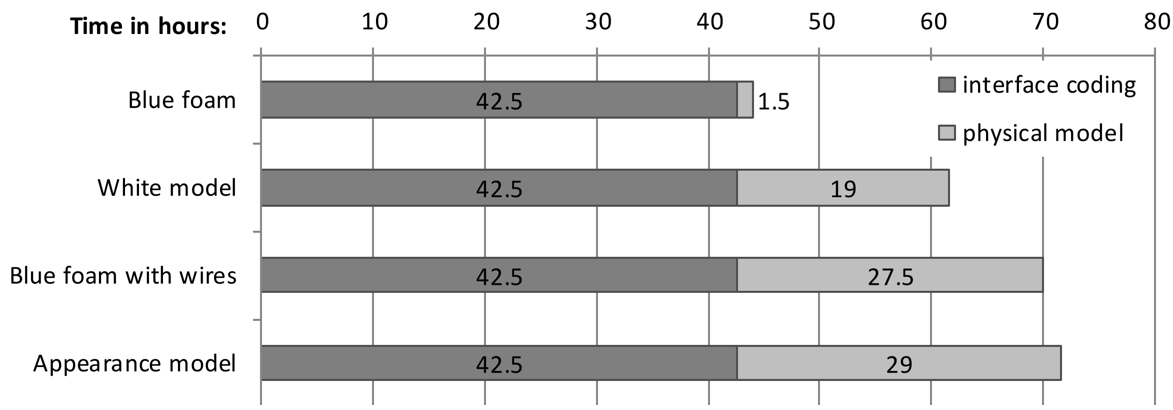


Figure 33: Time taken (in hours)

In addition to the time taken, the cost of bought-in materials was recorded. The specific bought-in materials are shown in Table 4.

Table 4: Materials used in each prototype

Blue foam	Blue foam
White model	Switches, dial, use of FDM machine
Blue foam with wires	Switches, dial, blue foam
Appearance model	Use of FDM machine, spray, printed graphics

From these two data sets an approximate cost for construction can be calculated for each prototype. The time costs were set at £40 per hour. It was assumed that the laptop, tablet, Arduino and IE4 were not bought specially for these prototypes and their associated costs are therefore not included. Figure 34 shows the costs of each prototype, with the 'appearance' prototype and 'blue foam with wires' prototype costing the most.

Figure 34: Breakdown of costs

5.5 Analysis of the prototypes

A description of the active and passive physicality aspects of the prototypes is shown in Table 5. From these descriptions each prototype was assigned an appropriate value of low, medium or high level of active and passive physicality, these levels are also shown in the table.

Table 5: Assessing the levels of active and passive physicality of the prototypes

Prototype	Passive physicality	Active physicality
Blue foam	Low This prototype looks approximate and feels light, buttons are obviously cardboard and not working.	Low Buttons are obviously intangible and the participant is speaking through their expected interactions which are being interpreted by the facilitator who is operating the Flash based interface.
White model	Mid This prototype looks reasonable with no distracting wires. The prototype can be held comfortably yet it is very obviously an early stage prototype.	Mid Interactions mimic the design intent satisfactorily directly operating the interface which is a reasonable approximation of the design intent.
Appearance model	High The prototype looks and feels very similar to the final product.	Low The interactions are not obvious as the participant does not use the tangible prototype to operate the interface; instead the interface is operated on a touch screen breaking the link between the tangible product and its interface.
Blue foam with wires	Low The prototype has tacked on switches and wires that are distractingly apparent in both the visual appearance and tangibility of this prototype.	High The prototype accurately mimics the way the final product feels when it is operated, both in the way the buttons work and the functionality of the interface.

Figure 35 is a graphical representation of the assigned physicality levels; the ‘appearance model’ and ‘blue foam with wires’ prototypes are high in one area of physicality at the expense of the other, whilst the ‘blue foam’ and ‘white model’ prototypes ‘balance’ both active and passive physicality.

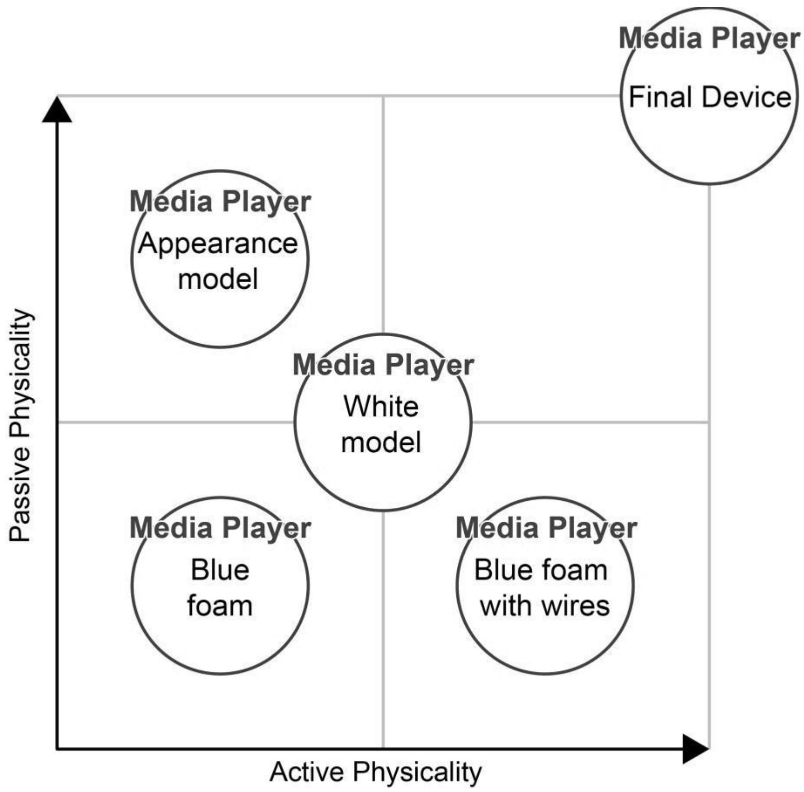


Figure 35: The resulting physicality of each of the prototypes

5.6 The User Study

5.6.1 Location

Participants were brought into a controlled environment (the PDR Insight Lab as shown in Figure 36), and the entire user trial was recorded on video. A moderator ran the study with an observer monitoring the study via a video link. The observer was intended to ensure continuity across the studies; this was deemed more suitable than introducing that person as a second moderator due to their level of experience with the prototypes and user testing methodologies. The moderator has conducted a number of similar studies before in a research and commercial context, and is therefore able to reflect on techniques with colleagues of similar experience.

Figure 36: The four video feeds showing the Insight Lab

5.6.2 Participants

A pilot study was first carried out with 4 participants from Cardiff Metropolitan University; these enabled a test run of the entire structure and the prototypes.

40 participants were recruited for the main study (eight per prototype (Goodwin, 2009)), two did not attend and three tests were rejected due to technical difficulties, so the total number included in the analysis is 35 (7 for each prototype and the final product). Each participant was randomly assigned to a prototype or the final product.

16 of the participants were female and 19 were male. Participants were screened in accordance with the target market identified by iRiver to be between 23 and 45 years old; recruited participants fell predominately into the <28 (49%) or 29-33 (34%) age groups. All listened to music on a dedicated player or mobile phone, and none had used the iRiver Spinn before.

5.6.3 The Experimental Protocol

The following structure was applied to every participant for each of the five prototypes trialled. These will be expanded in subsequent sections.

1. Participant is given an information sheet which gives details of the study.
2. Participant is asked to sign a consent form.
3. Participant fills in a demographic questionnaire.
4. Participant is introduced to the product.

5. Participant is given the opportunity to comment on the product.
6. Participant is given five tasks to carry out (detailed in Section 5.6.6). The moderator records whether the user experienced a success, minor problem, serious problem, or a catastrophe.
7. User-led exploration of main menu.
8. Semi-structured interview.
9. Participant is introduced to all the prototypes for an in-subject response and questionnaire.
10. Debrief

5.6.4 Pre-task collection of demographic information

This was conducted on-line through 'Survey Monkey', an online survey tool, and captured the details concerning: age, and details of their use of portable music products and mobile phones.

5.6.5 Description of the product

Once the questionnaire had been completed, the participant was introduced to the session and the iRiver Spinn through a verbal description, and some images of how it was envisaged to look (the same images as Figure 26 in on page 105). The product was described as a '*portable product to play and store music and which has the capability of playing other media types such as video and radio*'.

5.6.6 Tasks

Although completion rates were recorded, the tasks were primarily intended to introduce the participant to the product in a controlled manner. The tasks were not timed. Five tasks were chosen to introduce the participant sequentially to the product and no time constraint was imposed for the tasks. The tasks were:

Task 1: Turn on the music player.

Task 2: Play a specific track (named in the task)

Task 3: Adjust the volume of the track.

Task 4: Stop the track and navigate back to the main screen.

Task 5: Turn off the music player.

5.6.7 User-led exploration of the menu

Next, each participant was asked to scroll through the main menu titles and explain what they expected within each menu. This user-led exploration ensured each participant had the same knowledge of the features of the product for the semi-structured interview.

The main menu has ten options (Flash, Rec, Memo, Picture, Video, Music, Radio, Text, File and Set), which can be scrolled through with the wheel.

5.6.8 Semi-structured interview

A semi-structured interview sought to gain feedback about both the physical design and the users' interaction experience of the product. The semi-structured interview focused on four main areas, these were:

- Questions regarding the use of the product
- Feel and appropriateness of the dial
- Dial and digital interface
- Interface flow.

5.6.9 Comparison of all prototypes

Finally, users were introduced to all the prototypes, and asked to complete a questionnaire ranking the quality of feel, appearance and quality of interaction for each of the prototypes (6 point rating scale of 'positive' to 'negative'). This enabled the participants to directly compare prototypes and offer an opinion about their construction. The questionnaire also gave participants the option for a written comment after each of the three questions.

5.7 Data collected

As for user study one, a mixture of quantitative and qualitative data was gathered from the user study for analysis.

5.7.1 Quantitative Data

5.7.1.1 *Performance data*

The performance rating for each task was captured during the session by the moderator. The usability problems that caused users to struggle with tasks were also recorded for later analysis. All ratings and usability problems were subsequently reviewed using video capture data to ensure consistency with the above marking scheme.

5.7.1.2 *Comparison data for all prototypes*

Participants were asked to 'rate' (on a 6 point scale) each of the prototypes in three different areas (quality of feel, appearance, and quality of interaction) producing quantitative data. This questionnaire compared all prototypes to enable a quantitative look at a generally subjective area.

5.7.2 Qualitative Data

In addition to the performance ratings data, qualitative data on testing was gathered through:

- Comments and user interactions captured using video during the test.
- A semi-structured interview roundup at the end of the test.

All four cameras in the Insight Lab were used to capture each trial; this was transcribed and then coded thematically. Video clips were managed using the qualitative video analysis software *Observer XT* (Noldus, n.d.).

5.8 Comparative Analysis of the Prototypes in the User Study

5.8.1 Task completion rates

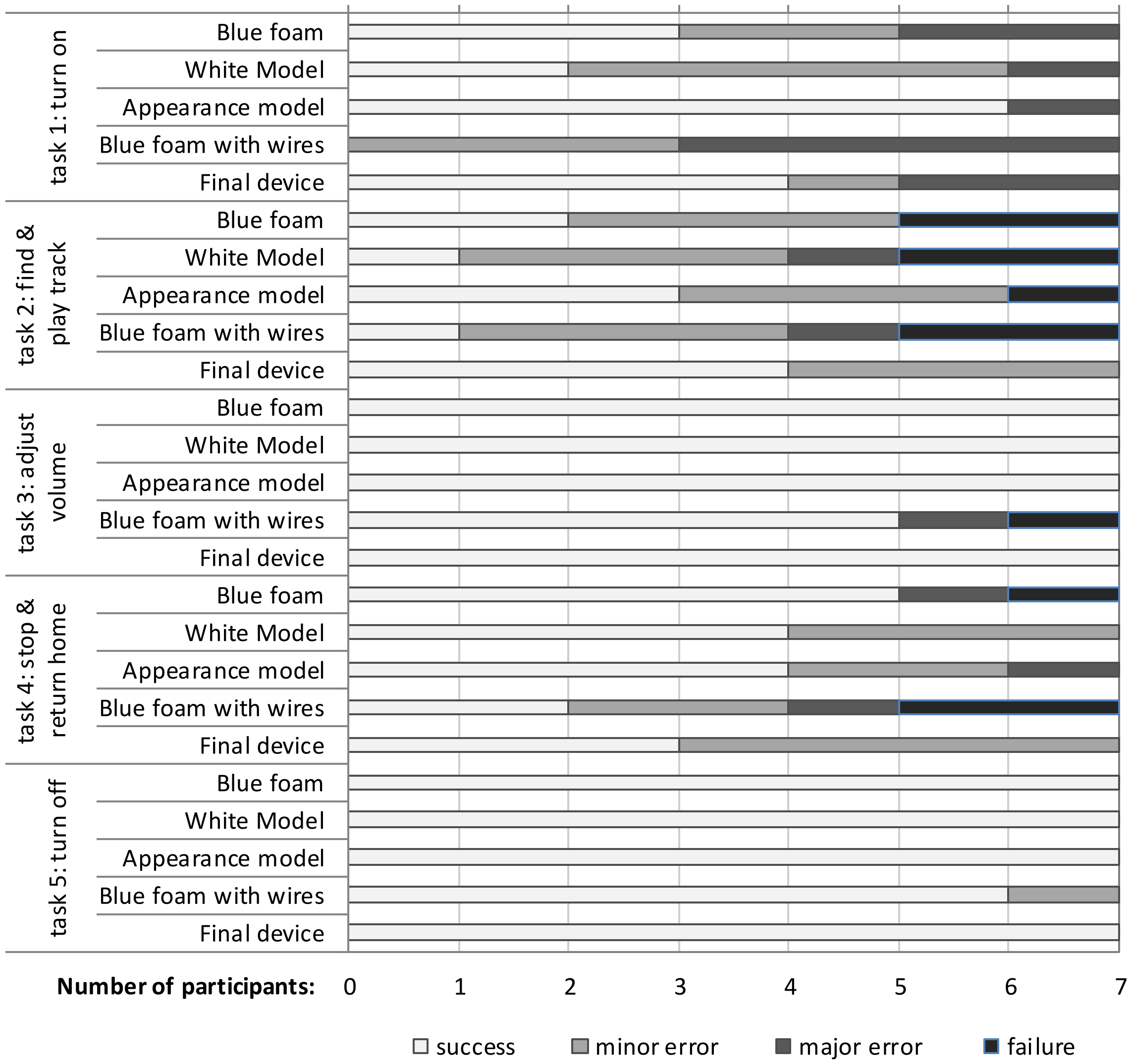


Figure 37: Task success rates

Figure 37 shows the task success rates for the study. As with Study One, the results of particular interest are those that differ across the prototypes. With Study Two each prototype can also be compared to the results of the final product. So, for example, during Task 1, the 'white prototype' had the highest successful completion rates of all the prototypes including the final product. Indeed, results from the 'white model' seem to gravitate largely towards the success or minor error end of the scale with only 4 participants experiencing major errors or a failure. Figure 37 was refined to focus solely on the number of major errors and failures experienced for each prototype for all tasks, the results are shown in Figure 38.

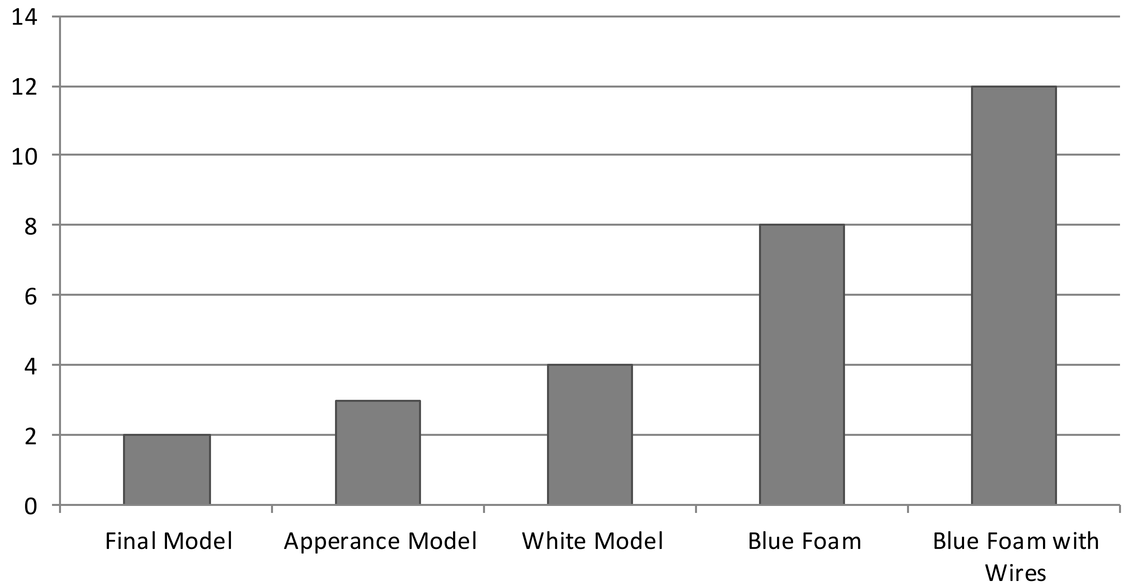


Figure 38: The total amount of major errors and failures for each prototype for all tasks

5.8.2 Discourse analysis

The analysis was performed by the moderator. Discourse analysis provided a framework to analyse the video footage of the tasks, menu exploration and semi-structured interview. The strength of this approach is that it gives the ability to structure the conversational feedback typical of this type of study in a rigorous manner. The video footage was reviewed with event logging software, from which comments and actions were coded into groups. The raw data from the tasks plus menu exploration and semi-structured interview can be found in Appendix 10. This also shows the comments grouped into the main comment codes.

For analysis, single comments were regarded as outliers and only comments made by two or more participants were compared. These comments were then reviewed and collated to form high-level design recommendations typical of a report from user trials (Molich, Jefferies, & Dumas, 2007).

Further recommendations could be drawn from the data produced by the studies but for the purpose of this study, only the comments that have emerged through the formal discourse analysis are included.

5.8.2.1 ***Design Recommendations***

The ten key comments from the discourse analysis are:

- DR 1. Help required from the facilitator
- DR 2. Difficulties in finding the required interaction
- DR 3. Tried other interactions
- DR 4. Pressed back to stop track playing
- DR 5. Tried turning dial to get to pause icon
- DR 6. Observation that it looks like a touch screen product
- DR 7. Like the 'Spinn' interaction
- DR 8. Long-winded interface
- DR 9. No unique selling point
- DR 10. Vertical menu navigation not obvious

The recommendations for developing the design are based on these ten comments and are described below (Table 6).

Table 6: Design recommendations from Study Two

DR 1. Help required from the Facilitator
<u>Description:</u> Users required help from the facilitator in order to progress with the task (note that help was only given when requested and absolutely necessary). *<u>Recommendation:</u> Re-consider the mental model of the interface and interactions. The other recommendations from this trial will demonstrate the areas to focus on.*
DR 2. Difficulties finding the required interaction
<u>Description:</u> Users have difficulties finding the appropriate interaction. *<u>Recommendation:</u> Consider how users think they should be interacting with the product, are there interactions that are not matching on-screen semantics?*
DR 3. Participant tried alternative interactions
<u>Description:</u> This is linked to DR7. Users are trying other forms of interaction to get the product to do what they want. This is not necessarily as important as DR7 because users are prepared to 'explore' their products and this should be supported by the interface with obvious routes 'back'. *<u>Recommendation:</u> As for DR7 plus provide routes back to encourage exploration.*

DR 4. Pressed back to stop

Description: Users are going 'back' in order to stop the music from playing. Often this is because they couldn't find the 'pause' or 'stop' button –this will be covered in DR10. The other reason they are doing this is because they are not sure which route to take, therefore when they press back and the music stops they are happy (because this is what the task asked of them). This is not a desirable feature because it would mean that users cannot explore their music library whilst listening to music.
Recommendation: Ensure that music continues to play when navigating away from the 'now playing' screen.

DR 5. Tried turning dial to get to pause

Description: Users have difficulties trying to work out how to get the song to pause. The intended interaction is for the users to press the dial, but users are turning the dial to scroll between the apparent 'active' icons on the screen.
Recommendation: This is a major usability error in the architecture and semantics of the interface. It could be addressed through a re-design of the graphical semantics so as not to 'suggest' there are many interactive components on the screen that can be scrolled between. Or the architecture could be re-structured to enable the user to scroll between active icons.

DR 6. Looks like a touch screen

Description: This is quite an interesting comment. It demonstrates the user's preconceptions about products that are now on the market. This comment was made during or after the trial, therefore the design of the interface suggests touch screen interactivity and the user expects this.
Recommendation: Decide on a direction, should this be touch screen or should the interaction design be developed to support non-touch screen engagement in a more meaningful way?

DR 7. Like Spinn

Description: Users give a positive reaction or comment to the 'Spinn' (dial) feature of the product.
Recommendation: Again this is a very interesting comment, it is very hard to give a firm recommendation because this comment should be taken within the context of all comments. On the face of it, the recommendation would be to choose the non-touch screen approach suggested above. However, having facilitated the trial, a much richer story is held in the data, one that needs to be explored further. Trial other forms of interaction to gain a better understanding of this comment.

DR 8. Long-winded

Description: This comment is one that can be used to clarify the above comment about the scroll wheel. Users were often getting frustrated when navigating through the music library, there is no way of 'fast scrolling' to speed up navigation. The second place this was bought up was on the home screen, users were not aware that the menu would repeat, many thought they needed to scroll back through to reach the other end of the menu.
Recommendation 1: Include a fast scrolling capability, e.g. if the dial is moved quickly then the library could change from individual names to A-Z listing.
Recommendation 2: Demonstrate that the menu repeats by showing the repeated menu item when the user is at the end of the list.

DR 9. No Unique Selling Point

Description: This is a fundamental part of the design that needs to be addressed from the core of the design idea. Users do not understand the value of this product. Much of the feedback focussed on the diversity of the functionality and users could not understand how some of the functions related to each other.

<u>*Recommendation:*</u> *Address the core values of the design –what is the key function of the product? Work on how to communicate the core values more effectively.*
DR 10. Vertical menu feels 'odd'
<u>*Description:*</u> *This comment concerns the shift from a horizontal scrolling menu (for small lists) to a vertical scrolling menu (for larger lists). Many users found this shift disconcerting, although they understood why it was needed and therefore would potentially be willing to get used to it.* <u>*Recommendation:*</u> *Investigate other means to represent the large amount of data needed in a music collection. For example, scrolling by A-Z, shifting the screen to vertical scrolling or a different architecture for finding tracks.*

Four prototypes were studied alongside the final product. The inclusion of the final product enables design recommendations to be drawn solely from the results of participants using the final product. As highlighted at the beginning of the chapter, reviews of the product were sought and two were selected; one from Engadget (Miller, 2009), and the second from Digital Trends (2009). These were treated as independent reviews and the results of this user study were compared to recommendations from the reviews. This comparison provides some confidence in the ability of the study to capture design recommendations in a reliable manner.

5.8.2.2 *The design recommendations obtained exclusively from the participants who used the final product*

The task success rate for the final product is shown in Figure 39 and the number of participants who experienced problems related to the design recommendations are shown in Figure 40.

- The mental model and semantics of the interface requires further exploration, some users had difficulties completing the tasks and expressed comments to this effect.
- Participants tried to use the Spinn dial to navigate between icons on the interface. This should be explored to determine if it is an appropriate way to navigate the more data rich screens. If it is not appropriate, the graphical design of the 'now playing' screen should be re-considered.
- Two participants thought the product was touch screen. The current interface is possibly producing an unclear message about whether this product is touch screen or not. Further exploration could be undertaken to determine if this is affecting the use of the product. *Note that the device is, in fact, touch screen. That only 2/7*

participants made this observation indicates the significance of this design recommendation.

- Four participants commented that the menu structure seemed long-winded. Exploration of how to convey the amount of data needed in a media library should be given reasonably high priority.
- Two participants questioned the main purpose of the product; this comment was received once they had reviewed the main functions. Consideration should be taken of the core values the product should convey.

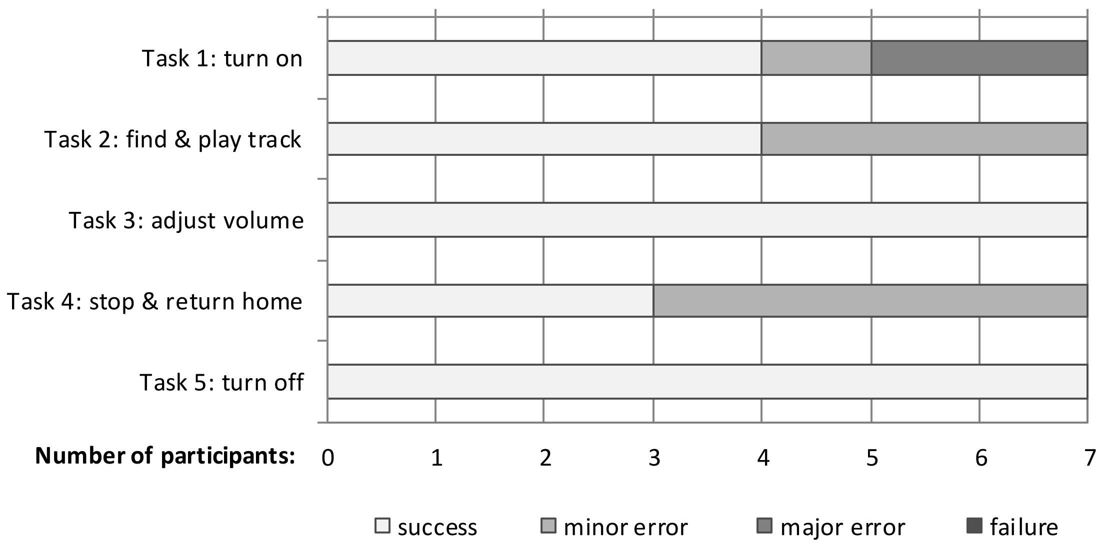


Figure 39: Task success rate for the final product

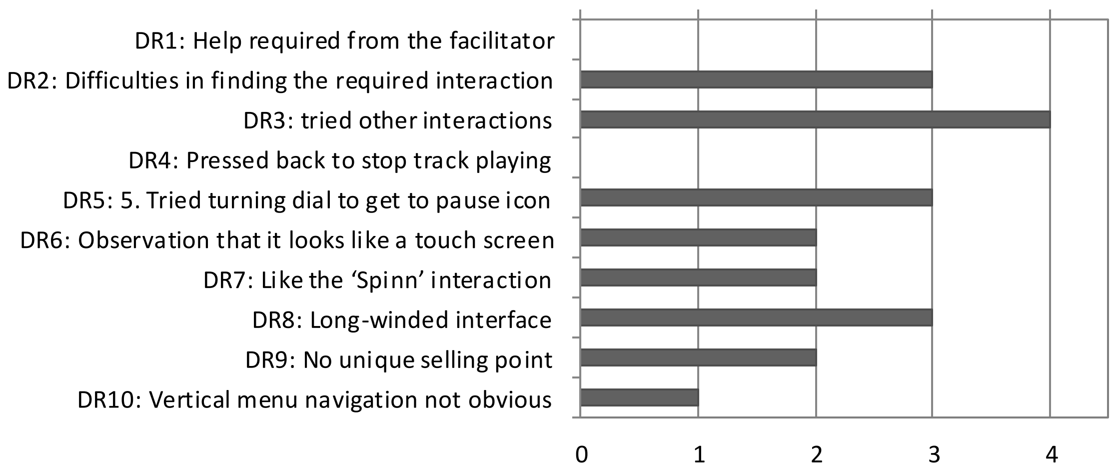


Figure 40: Number of participants who experienced the problems with the final product related to the design recommendations

Expert Review:

The Engadget review (Miller, 2009) is used here as an independent expert review. The following comments have been extracted from the Engadget review; each point is followed by the relevant design recommendation from the user study.

- "(The Spinn control is) *rather large for a hardware element in this day and age, and it's a joy to thumb around with. It doesn't spin "freely," instead giving a pretty solid click for every step of movement, and it doesn't take long to get the hang of breezing through menus with it."* **(DR 7)**
- *"The music player is just a little disappointing for a product of this stature. Everything is there […] but actually interfacing with the thing requires a lot of drilling into categories and then backing back out, something the player just seems particularly unsuited to do with its physical controls."* **(DR 2, 3, 5, 8)**
- *"It's great scrolling through long lists with the spin wheel, but that's just not enough."* **(DR 7)**
- *"The interface […] make*(s) *using the scroll wheel and the back button exclusively a bit of a chore."* **(DR 8)**
- *"Placement of the controls demands that you hold the product in a landscape orientation, and that you use both hands."* **(DR 10)**

5.8.3 The use of design recommendations in understanding the results

All of the comments from the expert review that relate to the design of the interface have been captured by the proposed design recommendations from the user study. In addition, the user trials provided some indications of why the issues were occurring.

Figure 41 shows the number of participants who made a comment about each of the ten design recommendations for each prototype. Not all of the prototypes captured all ten design recommendations, and when comparing the effectiveness of the prototype the detail of design recommendations above are not important to this research. However, the number of recommendations identified for each prototype in relation to the final product is of importance in this context (shown in black).

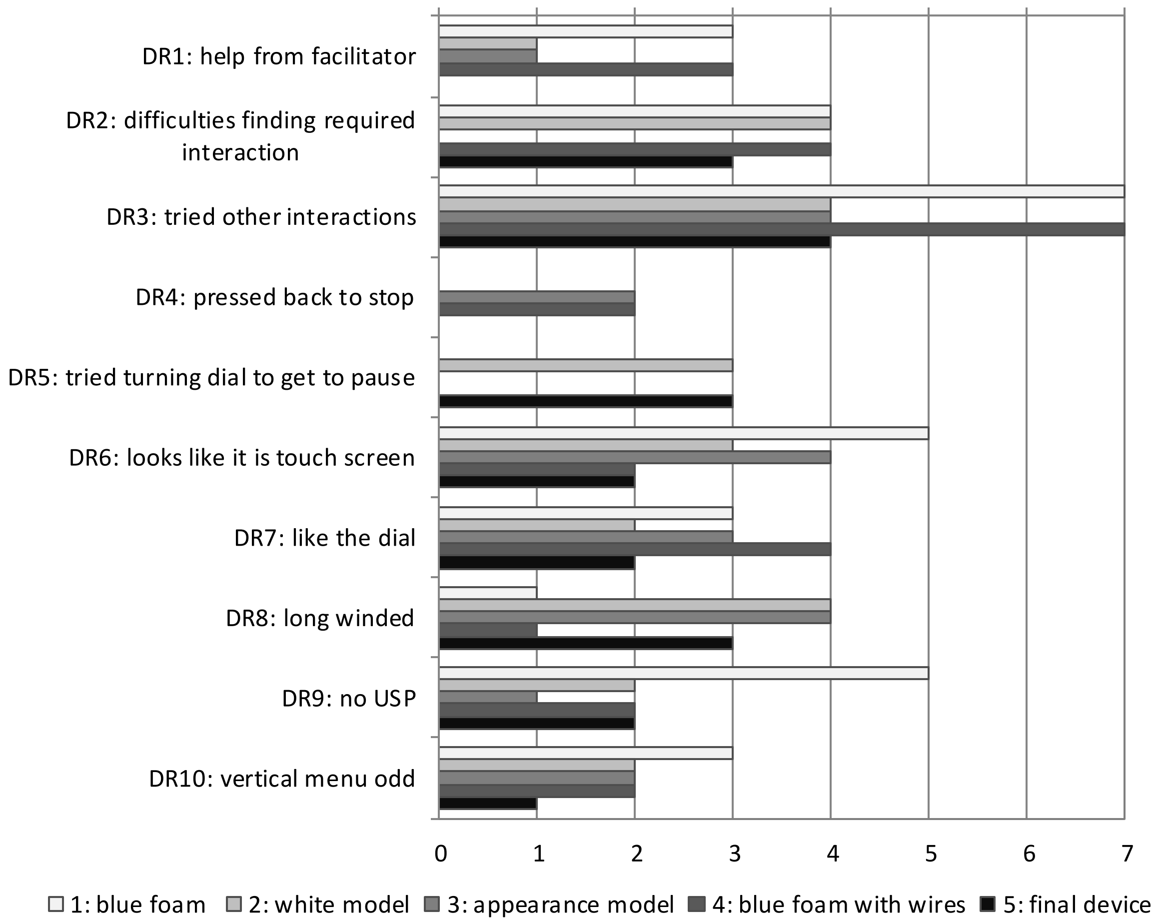


Figure 41: The ten key comments addressed by the design recommendations

5.8.4 Within subjects questionnaire

Figure 42 shows the results of the ranking exercise for which each of the participants were introduced to all the prototypes. The participants were asked to rate three elements of the prototypes; the 'quality of feel' and 'appearance' which aimed to prompt the participant to consider the passive physicality elements and the 'quality of interaction' which aimed to prompt the participant to consider active physicality. Although the terms 'quality of feel', 'appearance' and 'quality of interaction' cannot be directly described as active and passive physicality, it goes some way to enable a comparison to the assessment of physicality. The data from the prototype each participant used for the study was not included to eliminate any bias from familiarity with the prototype. Figure 42 shows participants consider the 'blue foam' prototype to have an equally low ranking for both elements, which supports the assessment of the prototype to be low in both active and passive physicality. Likewise, the 'appearance model' and 'blue foam with wires' are ranked in a similar way to the earlier assessment, with marked differences between active and passive physicality. The ‘white model’ however, produced interesting results, it was considered to have a higher ‘quality of interaction’ than the ‘blue foam with wires’, and a more marked difference between active and passive physicality than anticipated. It could be that the visual aspects of physicality are undervalued in the current definition of passive physicality, or that these questions are not adequate at obtaining participants views of active and passive physicality. This is beyond the scope of this study, but could be an interesting topic for further research. This exercise enabled participants to reflect on the prototypes themselves during the ranking exercise, and the comments made were also captured, these will be brought into the discussion.

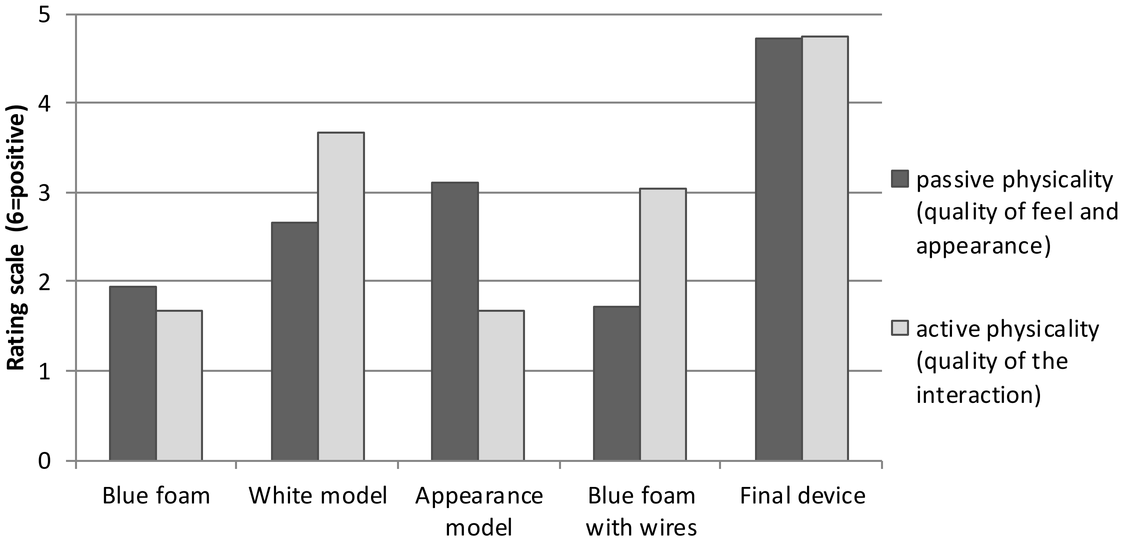


Figure 42: Data from the ranking exercise; comparing the physicality of the prototypes

5.9 Discussion

In Figure 41 and Figure 42, the 'white model' prototype appears to give feedback that is closest to the final iRiver product. These will be discussed along with other, more subtle, differences across the prototypes bringing in comments from the ranking exercise. Observations fall into two categories; 1. recommendations about the design and 2. obstructions caused by the prototype. Recommendations positively help identify how the design can be improved, whilst obstructions are caused by features of the prototype that hinder participants in giving meaningful feedback.

5.9.1 Recommendations about the design

5.9.1.1 *Physicality of the dial*

The 'white model' prototype was the only prototype that highlighted participants trying to turn the dial to get to the pause function (DR 5). This was also highlighted by the final product. The physicality of the dial itself could be the cause of this, for the 'white model' each rotation has a distinct 'click' which causes a reaction in the interface, but this was different from the final product. The dial on the 'blue foam with wires' prototype more accurately represented the final product because its dial had more subtle feedback. This suggests that there must be something else about the prototype that causes the participant to miss feedback for this design recommendation. Several users made comments about the wires of the 'blue foam with wires' prototype being "very distracting" and looking "messier" than the other prototypes, this 'messier' appearance could possibly be the cause of this.

5.9.1.2 *Information architecture*

The feedback that the interface was longwinded was a common comment from participants of the trial with the final product. The 'white model' and 'appearance model' model both produced the same feedback. The 'blue foam' prototype was not able to elucidate this, possibly because the participant was not directly manipulating the prototype and therefore not creating the direct mental link between the physical and digital. Comments which indicated this included; *"I did not like the fact that I couldn't control the product (interface) from the model"*. Meanwhile, the 'blue foam with wires' prototype also produced few comments about this possibly because the novelty of the prototype itself suppressed the participant's potential frustration with the navigation of the interface. The 'white model' seems to give a very direct feel between the interface and interaction, mimicking the final product well. The ‘appearance model’ forced the participant to have to continually press the scroll button to navigate the interface, highlighting the sheer number of button presses required to navigate the interface. As one participant observed: *“[It is] very tedious going through all the songs like this”*.

5.9.2 Obstructions caused by the prototypes

5.9.2.1 *Modelling physical interfaces on a touch screen*

The ‘appearance model’ used a touch screen for the interactive element of the prototype. This prototype gave participants the least difficulties in finding the interactions. As a result of the need to represent all the buttons on a touch-screen this prototype clearly indicated where interactions were, even when they were on the side of the product (as shown in Figure 43). This made the interactions more obvious for those using this prototype than they would otherwise have been. Paradoxically, given the number of issues users had with the real product, the ease of use of the interface on the touchscreen reduces the effectiveness of the prototype.

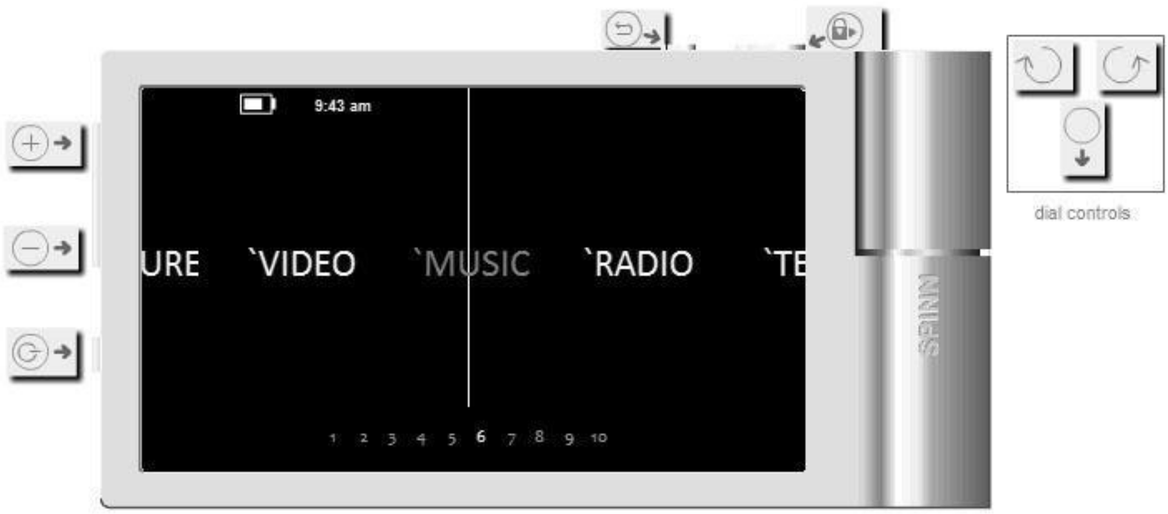


Figure 43: The on-screen prototype

5.9.2.2 *Obstacles to the participants understanding the prototype*

Figure 41 shows the ‘blue foam’ and ‘blue foam with wires’ prototypes forced participants to ask for the most help from the facilitator. The ‘blue foam’ model requires the participant to fully engage with the ‘speak aloud protocol’ because the buttons provide no active feedback. The participant therefore has to wait for the moderator to operate the interface. In contrast, the ‘blue foam with wires’ prototype allows the participant to operate it independently, but it may be the appearance of the wires that seems to be the biggest barrier to acceptance. It may also be that techniques which require the participant to understand the *way* in which the prototype works are not suited for this type of early stage trial.

5.9.3 Overview of the four prototypes

5.9.3.1 ***The 'white model' prototype***

The real-time nature and simplicity of this prototype seem to be the important factors in making this prototype the most effective of the prototypes. Participants were able to operate and receive immediate feedback from the interface without an overly complicated looking prototype or altering the scale and form of the model. For example, one participant said: *"I felt very little difference in terms of the final version and 'white model' for the quality of interaction – [the] 'white model' had a few blips but nothing that is stopping me using the product successfully."* and another said: *"the addition of working buttons on the prototypes increases the quality of the feel, as the ways in which interaction occurs can be more readily envisioned".*

5.9.3.2 ***The 'blue foam' prototype***

This prototype used the 'speak out loud' protocol for participants to engage with the interface. Results show that this prototype was less effective at enabling participants to build a mental model of the product resulting in reduced effectiveness of the comments received; *"The colour, weight, size and cable connections play a big part of my initial interaction with a product, for this reason the 'blue foam' compared to the final unit was clearly a visual aid as opposed to actual real product comparison."*

5.9.3.3 ***The 'blue foam with wires' prototype***

Participants required more assistance using this prototype. This was a surprise from the prototype with the highest fidelity interactions. Participants seemed to be affected by the wires and appearance of this prototype; *"The model with blue foam & wires looks messier than the blue foam model but it looks a little bit more functional than the model with blue foam alone."*

5.9.3.4 ***The 'appearance model' prototype***

This prototype used a touch screen to convey the interactions of the prototype. Participants did not identify as many usability errors and had the most differing performance in relation to the final product. This outcome supports Gill *et al.*'s study in which it was proposed that interactions are easier for a participant to identify on a screen

(Gill, et al., 2008); "*Although the silver model (appearance model) looked more like the final version, I did not like the fact that I couldn't control the product from the model, and I didn't think having the model alone, without much interaction, was very worthwhile*."

5.10 Conclusion

The four prototypes trialled in this study explored different aspects of active and passive physicality. The results show that both active and passive physicality are important considerations for early stage user feedback; but it is an even proportion of these that produces the most effective prototypes, as seen in the 'white model' and 'blue foam' prototypes. In this situation, an 'effective' prototype is one which elicits feedback related to the intended design to enable the next iteration of the design to take place. Resources should not be used exclusively to ensure the interactions of the prototype are high fidelity (active physicality) if it severely impacts the ways the prototype looks, or can be held by the user (passive physicality). Likewise, resources spent creating a prototype that looks very close to a final product are not effective if interactions are not well supported.

The 'white model' and 'blue foam' prototype provided the most accurate data compared to the user experience of the real product. Both the 'white model' and 'blue foam' prototype were of balanced physicality. The 'blue foam with wires' was very strong on active physicality to the detriment of passive physicality whilst the 'appearance model' model was very high on passive physicality but low on active physicality. This suggests that it is those prototypes that are well balanced that are the most effective in this study. Since they are also cheaper they represent strong value for money.

The prototype has long been accepted as a valuable approach to creating insightful design outputs. However, for interactive products that have both a physical and digital form, visual interface fidelity alone is clearly not enough to fully conceive the complete prototype and ensure it will accurately fulfil its purpose. Whilst visual and dimensional fidelity is very much the staple of prototyping, physical fidelity clearly has a role in creating an effective prototype. This study indicates that for interactive prototyping, 'physicality' needs to be an even combination of both active and passive physicality.

Chapter 6. Discussion and Conclusions

This chapter will explore the hypothesis of balancing passive and active physicality by re-examining the Home Phone prototypes discussed in the Contextual Study (Section 2.4.5), the conceptual Photo Management prototypes of Study One (Chapter 4), and the Media Player prototypes of Study Two (Chapter 5). This provides eleven prototypes in total (not including the final products); an overview of these prototypes can be found in Table 7 below.

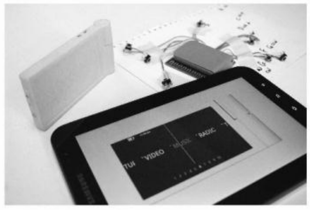
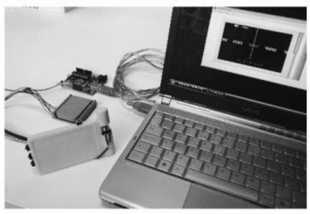
As a result of Study One it was proposed that the physicality of a prototype can be considered on two levels; that of active and passive physicality where; passive physicality is proposed to be the perceived affordance based on the tangibility and aesthetics of the prototype, and active physicality is proposed to be the tangible and visual experience of interacting with the prototype.

The results of Study One and Study Two have enabled an in depth exploration of the constructs of active and passive physicality which has exposed a number of areas which need further development. The first section of this chapter will develop the construct for active and passive physicality so that it can be effectively applied to all eleven prototypes. The subsequent sections re-visit the prototypes of each study to assess them in terms of the updated construct of active and passive physicality. This new understanding of the prototypes is subsequently used to re-examine the data of each trial, through a 'critical reflection', to determine if physicality can provide an insight into the results obtained. Once the prototypes of all three studies are re-examined, all eleven prototypes are compared in order to test the hypothesis of active and passive physicality.

The following paper, which discusses the 11 prototypes in relation to active and passive physicality, has been accepted for publication in the Journal of Design Research and can be found in Appendix 11:

> Hare, Gill, Loudon & Lewis (2014), **Active and passive physicality: making the most of low fidelity physical interactive prototypes**. Accepted for publication in the Journal of Design Research. Inderscience.

Table 7: Overview of the eleven prototypes plus the final products

	Contextual study: **Mobile Home Phone**	Study One: **Conceptual Photo Management Product**	Study Two: **Media Player**
Prototypes	 Software only		 Blue foam
	 Sketch	 Low fidelity	 Appearance model
	 Flat face	 Mid fidelity	 White model
	 High fidelity	 High fidelity	 Blue foam with wires
	 Final device		 Final device

6.1 Refinement of the Construct of Active and Passive Physicality

The prototypes created for this thesis have enabled an in depth exploration of the constructs of active and passive physicality. In addition, the three articles specifically discussing the physicality of interactive prototype have enabled further discussion of the constructs of physicality with other experts in the field (two conference papers and a journal paper in press). Both of these activities have identified a number of areas which need further development including:

1. The terminology of the description of active and passive physicality
2. Where the distinction between active and passive physicality lies
3. The relationship of physicality to Physigrams.

6.1.1 The terminology of the description of active and passive physicality

Specifically, the term 'aesthetics' within the description of passive physicality did not convey what was intended and this has been modified to 'visual appearance'. This confusion seems to be related to what different disciplines understand by the term 'aesthetics'. The original inclusion of the term was intended to convey the visual qualities of the prototype, however, the term is also used to describe the 'critical reflection of art, culture and nature' (Kelly M. , 1998, p. ix).

Thus, the updated description of passive physicality is the **perceived affordance mainly based on the visual appearance and tangibility of the prototype.**

For active physicality, the wording 'tangible and visual experience' was found to be confusing and was therefore updated to the 'perceptible experience'. The original description was found to limit the application of active physicality, ignoring additional feedback which does not happen on screen such as lights illuminating, a mechanism engaging or vibro-tactile feedback.

Thus, the updated description of active physicality is **perceptible experience of interacting with the prototype.**

6.1.2 Where the distinction between active and passive physicality lies

When considering what happens at the point of interaction, the boundary between active and passive physicality became confusing using the original definition. For example, what type of physicality is involved at the point of pressing a button? Active physicality is initiated through the button press, but is what is felt active physicality or passive physicality because it can be felt whilst the prototype is 'unplugged'? Further problems arise when considering products which include vibro-tactile feedback. Vibro-tactile feedback could be interpreted as passive physicality because it is the feedback of interactions, but it requires the prototype to be 'switched-on' and therefore relates to active physicality.

Therefore the boundary between active and passive physicality has been defined as the '*point at which manipulation of the product occurs which requires a system response or mechanical action (or both)*'. For example, the sense of touch is used to determine whether buttons fall in a 'natural' location (passive physicality) but if interaction with those buttons occurs, to determine what they do and how they feel, this now falls under active physicality. If those actions are intended to initiate further actions, for example, changing a screen element, this should be considered alongside its tactile feedback. An interaction which does not comprise all of its intended actions will have lower active physicality than one that does. Take the appearance model of Study Two for example, the switches are not connected, but they will deform and feel like they should but interaction will not result in any feedback beyond the tactile.

Interaction with the buttons can now be determined to fall entirely under active physicality despite the fact some of this detail can be felt while the prototype is 'unplugged'. This relationship between the product 'unplugged', as identified in the Contextual Study of the Physigrams, is discussed in the next section.

For the vibro-tactile switch example given above, the resultant feedback of interaction requires the system to respond to the action and therefore entirely falls under active physicality. In addition, any perceptible feedback that comes from interaction will fall under active physicality, even if the prototype is 'unplugged' and the feedback is not the intended vibro-tactile feedback which would have occurred if the prototype was 'switched-on'.

6.1.3 The Relationship of Physicality to Physigrams

Despite the major influence of the Physigram study in creating the constructs of active and passive physicality, the last section has identified that considering the product 'unplugged' does not fully explain the difference between active and passive physicality.

The problem with considering the product 'unplugged' in terms of active and passive physicality lies in the relationship between the interaction and its feedback. The intention of the Physigram study was to focus solely on the physical feedback of interactions, whereas the intention of this thesis is to consider the effectiveness of the entire interactive prototype. Since interaction is meaningless without feedback in this context, considering the prototype 'unplugged' in the same sense as Physigrams is not useful. However, the example of considering the prototype 'unplugged' is certainly an effective way to start the description of the difference between active and passive physicality.

6.1.4 Hypothesis of Active and Passive Physicality

This thesis proposes that the physicality of a prototype can be considered on two levels; that of active and passive physicality where; passive physicality is the **perceived affordance based on the visual appearance and tangibility of the prototype,** and active physicality is the **perceptible experience of interacting with the prototype.**

It hypothesises that both active and passive physicality can be considered on a scale of low to high and that prototypes which fall below certain levels of either active or passive physicality in relation to the design intent are least effective and prototypes that balance active and passive physicality equally are the most effective. In this situation an 'effective' prototype is one which elicits feedback related to the intended design to enable the next iteration of the design to take place.

The proposal that active and passive physicality should be 'balanced' recognises that many prototyping construction techniques require a compromise of some kind. For example, the use of electronics within a prototype necessitates components and power requirements which could impact the size of the prototype and the demand for a highly realistic prototype could impact the way in which the prototype can be interacted with. In these scenarios the resultant physicality of the prototype is affected even though its fidelity is not necessarily altered. Without an understanding of this affect any prototype created could be limited in its effectiveness.

Figure 44 shows the hypothesis of active and passive physicality in graphical format. The two axes show the extent of active and passive phsycialty with the orgin being 'no physiclaity' and the 'highest' physiclaity being equivalent to the final design intent. The areas below the dashed lines show that there it a point at which physiclaity can be considered too low for an effective prototype. The diagonal depicts where prototype considered to have 'balanced physicilaty' should fall on the graph.

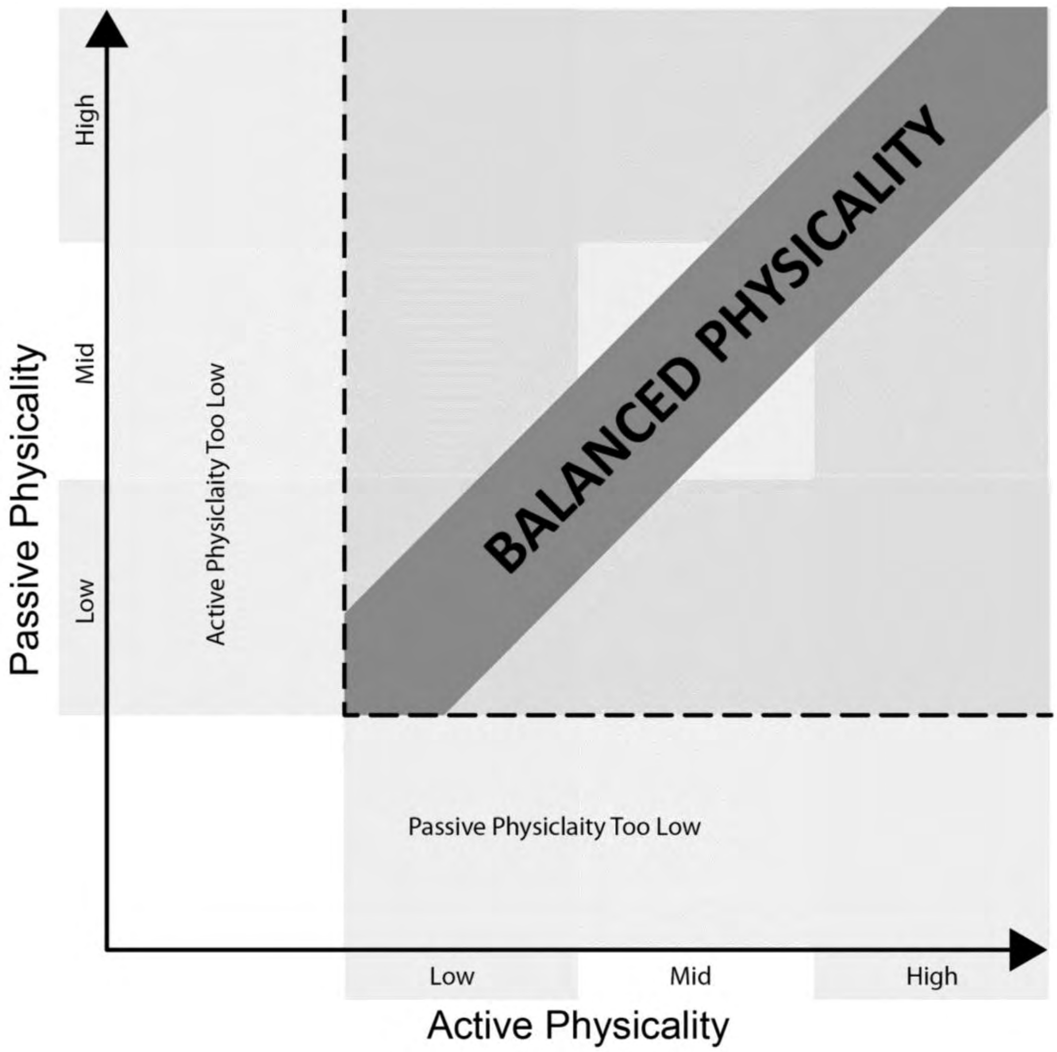


Figure 44: The framework of active and passive physicality

6.2 Contextual Study: Mobile Home Phone

This study was introduced in Section 2.4.5. The study explored the effect of fidelity on a series of prototypes of a Mobile Home Phone. These prototypes are shown in Figure 45 (redrawn here for convenience). The objective of the study was to determine if the results of a user trial with a physical interactive prototype were more similar to the final product than a software-only prototype, and the subsequent level of fidelity required of the physical interactive prototype. The technique used to construct the prototypes was determined by fidelity levels.

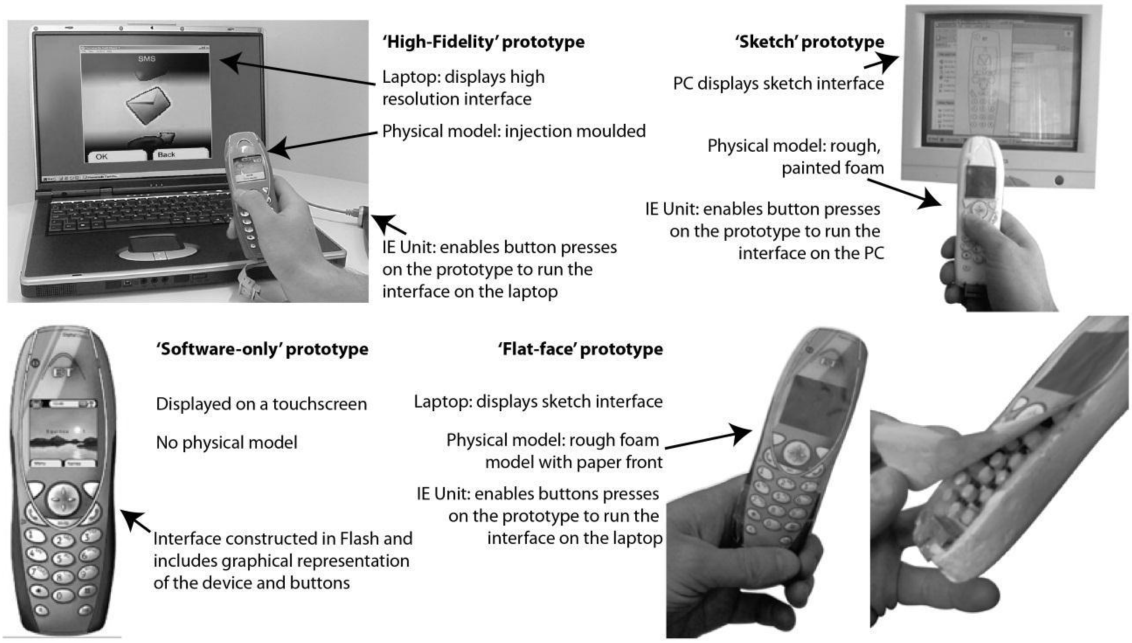


Figure 45: The four prototypes used in the Home Phone Study

6.2.1 Assessing physicality

The physicality levels of all the prototypes are shown graphically in Figure 46. The tangible and visual qualities of the physical model of the '**high-fidelity**' prototype are very similar to the final product, with the weight and appearance of the wires (connecting to the IE unit) being the only compromises (high passive physicality). Upon interacting with the prototype, the buttons have the same feel as the final product with the onscreen graphics performing to a high-fidelity, albeit on a remote screen (high active physicality).

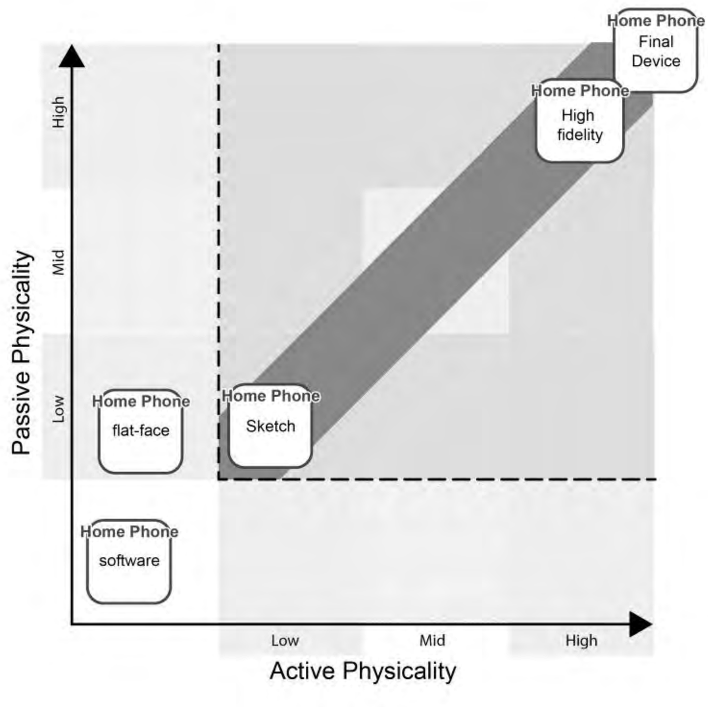


Figure 46: The physicality levels of the Mobile Home Phone prototypes

The visual appearance of the '**sketch**' prototype is very crude but the tangible aspects of scale, form and button location are a good approximation of the final design (low passive physicality). Upon interaction, the buttons have the similar feel as the final product and the onscreen graphics are very simple in appearance. The structure of the interface is identical to the high-fidelity prototype although the functionality was reduced (low active physicality).

The scale and form of the '**flat-face**' prototype are restricted due to the front being removed, the printed visual appearance is reasonable and the buttons appear to be in a good approximate location (passive physicality is marginally lower than the 'sketch' prototype). Yet upon interaction it becomes apparent that the 'hit' area of the buttons differs from what is visible on the surface. In addition, the quality of the physical feedback of the buttons was reduced by the paper. The interface was identical to the sketch prototype (active physicality is significantly lower than the 'sketch' prototype).

There is no tangible model for the '**software-only**' prototype; therefore the only concession to passive physicality is a two-dimensional graphical presentation of the design resulting in extremely low level of passive physicality. The interface is identical to the 'high-fidelity' prototype, yet interaction with the interface is vastly different to the final product with no tactile feedback of the physical device or buttons. This marks a very interesting attribute of active physicality, the lack of a physical device to hold and manipulate significantly lowers active physicality levels despite the onscreen interface being considered 'high-fidelity'.

6.2.2 Overview of Results

User trials were conducted utilising the four prototypes and the final product. Users were asked to complete six tasks and the success rate of each task was recorded.

The 'high-fidelity' prototype produced similar results to the final product, significantly outperforming the 'software-only' prototype. The 'sketch' prototype was found to perform similarly to the final product. The performance of the 'flat-face' prototype however, was significantly reduced. It appeared that the flat face of the prototype did not replicate the true physicality of the product sufficiently, and that this resulted in more user errors, slower performance times, and worse performance ratings.

The initial results of this study suggested that it is not the level of fidelity that is important, but rather the considerations of tangibility and physicality. Specifically for the 'flat-face' prototype, the study proposed that there was something which was lacking in the physicality of that prototype which prevented it from being effective. The critical reflection looks to the constructs of active and passive physicality to provide a means of better understanding these results.

6.2.3 Critical Reflection

When the hypothesis of active and passive physicality is considered, it becomes apparent that the 'software-only' prototype does not have any form of passive physicality and (surprisingly) little active physicality. What is surprising in this case is that despite the interface being identical to the high-fidelity prototype, there is a considerable difference in way in which the interface is operated. On the software-only prototype, there is no physical model or buttons with which to operate the interface, and this has a significant impact on the active physicality level because the user cannot tangibly feel the model or interaction.

In this trial, the 'sketch' prototype, although low fidelity, implements enough active and passive physicality for the user to understand the design on a similar level to the 'high-fidelity' prototype. This reveals a significant saving in time and expense in terms of constructing a prototype, in addition to being able to construct this type of prototype earlier in the design process enabling more iterations of the design.

On initial appraisal, the 'flat-face' prototype appeared as though it would produce effective results because the only difference between this and the sketch prototype is the paper covering the buttons. But when notions of active and passive physicality are applied, it becomes apparent that passive physicality is very low in comparison to the design intent. Feedback of the interaction is poor since the participant cannot determine exactly where the 'hit area' is underneath the paper, resulting in unsatisfactory feedback upon interaction (active physicality). In addition, the interactions are not transparent enough for the user to understand how to operate the prototype, in other words, the appearance and tangibility of the prototype suggest there is little the participant can do with the prototype (perceived affordance resulting from passive physicality). This prototype has been a really interesting case study because it marginally challenges the boundaries of an acceptable low fidelity interactive prototype.

6.3 Study One: Conceptual Photo Management Product

Study One is covered in detail in Chapter 4. This study sought to explore the role of physicality on three early-stage interactive prototypes of a conceptual photo management product, shown in Figure 47 (redrawn here for convenience). The technique used to construct the three prototypes was determined by allocating time limits during construction.

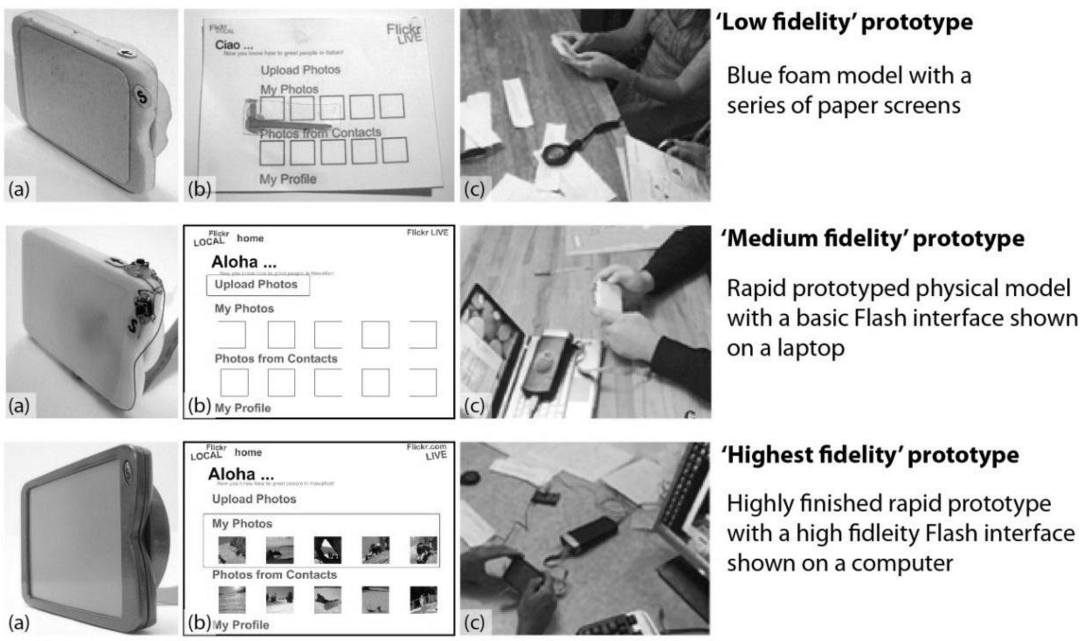


Figure 47: The three prototypes used for Study One

6.3.1 Assessing physicality

The levels of physicality are shown in Figure 48. Although the physical form is relatively accurate for the **low-fidelity** prototype, it feels very lightweight; interactions are clearly depicted but perceptibly non-functional, therefore this prototype has low passive physicality levels. Interaction relies on the participant pressing cardboard buttons and talking through their actions with the facilitator interpreting this by adjusting the paper screens. Although buttons are accurately located on the prototype, there is little tactile feedback of the buttons and delayed visual feedback of the interface; therefore active physicality is considered to be very low for this prototype.

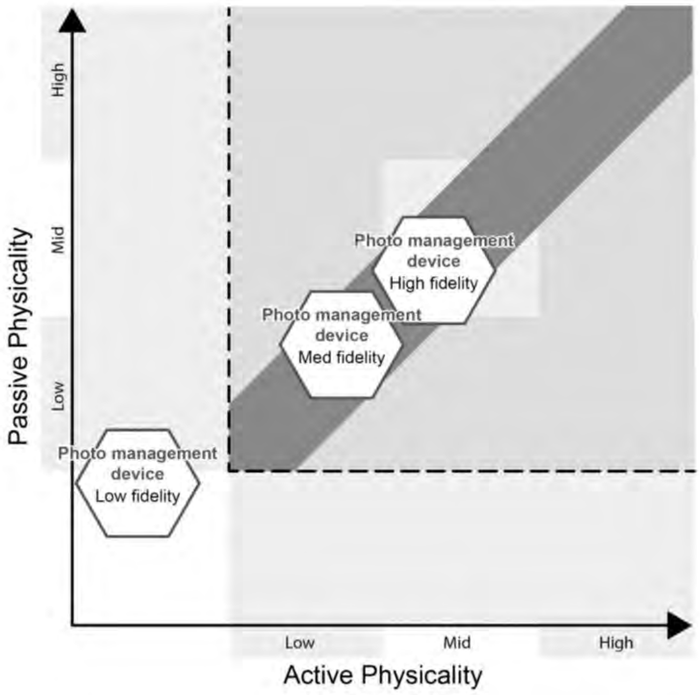


Figure 48: Assessment of physicality for the conceptual Photo Management prototypes of Study One

The physical form factor of the **medium-fidelity** prototype is relatively accurate; the unfinished form and tacked on buttons inform the user that interaction is possible, but it does not visually reflect the final product. Therefore the passive physicality of this prototype is low but still higher than the 'low-fidelity' model. The dial gives haptic feedback but this is not representative of the intended design; this dial feels 'clunky' and cannot rotate 360 degrees whereas the intended design fully rotates giving more subtle haptic feedback. Visual feedback of the interaction is immediate and the interface is functionally accurate but screen animations are not as refined as the intended design, therefore active physicality is higher than the 'low-fidelity' model.

The **high-fidelity** prototype was constructed and finished to accurately represent the intended design visually and tangibly (high passive physicality). It could be further improved by ensuring the weight of the prototype is more accurate. The interactions of the product reflect the intended design well with the dial providing full rotation with subtle haptic feedback and the interface includes good visual feedback (high active physicality).

6.3.2 Overview of Results

Users were asked to perform five tasks on the prototypes, task success rate and discourse analysis was performed on the resulting data. Initial analysis showed that task success rate did not differ significantly across the prototypes, although the results suggested that the greatest difficulty for users of the 'low-fidelity' prototype was identifying the correct interaction; whilst users of the 'medium' and 'high' fidelity prototypes had more problems creating a 'mental model' of the interface. Discourse analysis revealed that the 'medium' and 'high' fidelity prototypes were more effective at eliciting useful user comments than the 'low-fidelity' prototype. The next section looks to the constructs of active and passive physicality to provide a means of better understanding these results.

6.3.3 Critical Reflection

When the hypothesis of active and passive physicality is considered, it can be seen that, despite being of very different fidelity, the physicality of the 'medium' and 'high' fidelity prototypes is relatively similar. Therefore, in terms of physicality, very little has been added to the 'high-fidelity' prototype despite the additional time spent creating the prototype. The 'low-fidelity' prototype, however, is very low in active physicality due to the lack of haptic and visual feedback of the prototype. Despite setting out to assess the effect of physicality, when the notions of active and passive physicality are applied, the prototypes used in this study are fairly similar. This explains why the results of this study were inconclusive.

6.4 Study Two: Media Player

Study Two is covered in detail in Chapter 5. The technique used to construct the four prototypes was determined by the definition of active and passive physicality proposed as a result of Study One. The prototypes are shown in Figure 49 (redrawn here for convenience). The intention was to include the four permutations of active and passive physicality levels as demonstrated by each quadrant of Figure 28 on page 107. An existing media player product was chosen for retro-prototyping providing a datum product for comparison.

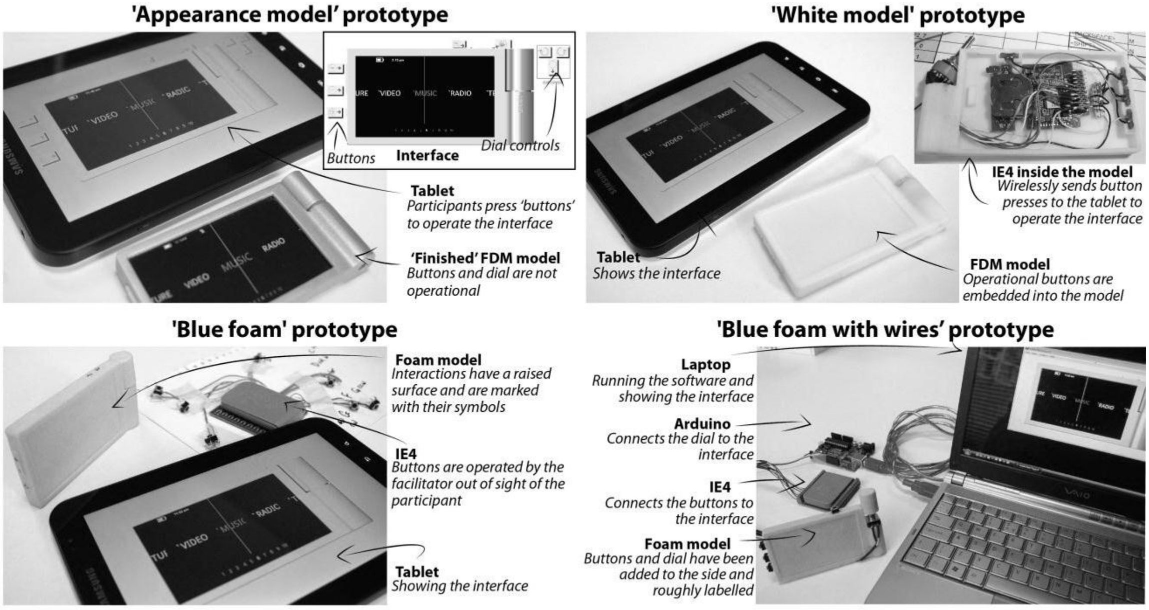


Figure 49: The four prototypes used for Study Two

6.4.1 Assessing Physicality

The levels of physicality are shown in Figure 50. The tangible and visual qualities of the **blue foam** prototype are accurate but low-fidelity and interactions are clearly not functional resulting in low passive physicality. Interaction is based on the 'speak out loud' protocol and operated by the facilitator, buttons are cardboard but the dial does rotate, therefore active physicality is also low.

The scale and form of the **white model** was slightly modified to incorporate the electronics needed. Its weight and the location of buttons are a good approximation of the intended design, and therefore this prototype has higher passive physicality than the 'blue foam' prototype. Upon interaction, the haptic and visual feedback is a good approximation of the final product, again resulting in higher levels of active physicality in relation to the 'blue foam' prototype.

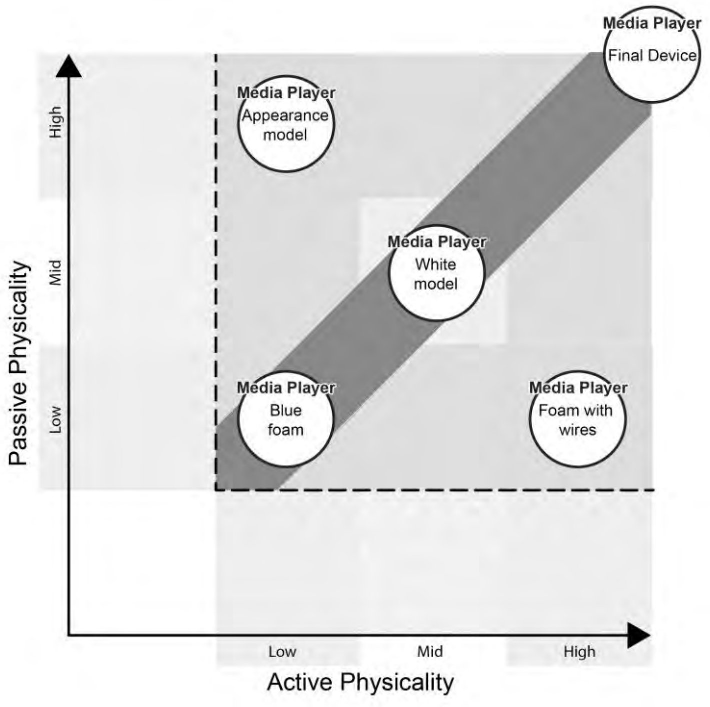


Figure 50: Assessment of physicality for the Media Player prototypes of Study Two

The **appearance model** has no electronics embedded in the physical prototype, therefore the scale, form, finish, weight and button location are a very accurate representation of the final product (high passive physicality). However, the absence of electronics means there is no feedback from the buttons or dial; and the interface is operated on a touch screen separate to the physical prototype resulting in low active physicality. This prototype really emphasises the distinction between active and passive physicality because the buttons on the model have good haptic feedback, yet this does not raise the assigned active physicality level because they do not function.

The **blue foam model with wires** has an approximate physical model that is clearly modified to accommodate the switches and dial; the wires are very apparent and visually impact the prototype resulting in low passive physicality. Upon interaction, the visual and haptic feedback of the buttons and dial accurately represent the final product making this prototype high in active physicality.

6.4.2 Overview of Results

Users were asked to perform five tasks before commenting on the main menu options, this ensured each participant had the same knowledge of the product for a semi-structured interview. The data was analysed to elicit design recommendations for each prototype and these were compared to the final product.

Participants using the 'white model' gave good feedback indicative of the final product. Results of the 'blue foam' prototype show that this prototype was less effective at enabling participants to build a mental model of the product resulting in less informative comments. Participants struggled to relate the action they were performing on the physical model to what was happening onscreen (active physicality).

Participants using the 'foam model with wires' prototype required more assistance using the prototype. This was a surprise given that this prototype had the highest active physicality levels. Participants seemed to be affected by the wires and appearance of this prototype (its passive physicality) resulting in less meaningful comments. The 'appearance prototype' had the weakest performance; although some interesting comments were received. The comments elicited by this prototype did not accurately reflect those of the final product. The critical reflection looks to the constructs of active and passive physicality to provide a means of better understanding these results.

6.4.3 Critical Reflection

This study was constructed to explore active and passive physicality by deliberately creating prototypes with different ratios of active and passive physicality. In terms of eliciting meaningful data, the 'white model' and 'blue foam' prototypes were more successful, with the 'white model' prototype receiving more useful feedback than the 'blue foam' prototype; both these prototypes balance the levels of active and passive physicality.

Active physicality for the 'blue foam with wires' was close to the design intent to the detriment of passive physicality, which fell someway short of the design intent. The 'appearance model' reversed this, with passive physicality close to the design intent at the expense of active physicality, which was not. Both these prototypes are typical at an early stage in the design process if a 'higher' fidelity prototype is required.

The study shows that both active and passive physicality are important considerations for early stage user feedback; but it is an even balance of these that produces the most effective prototypes.

6.5 Comparing all the Prototypes

The critical reflection of each of the studies presented supports the hypothesis of passive and active physicality, providing a framework to understanding sometimes unexpected results.

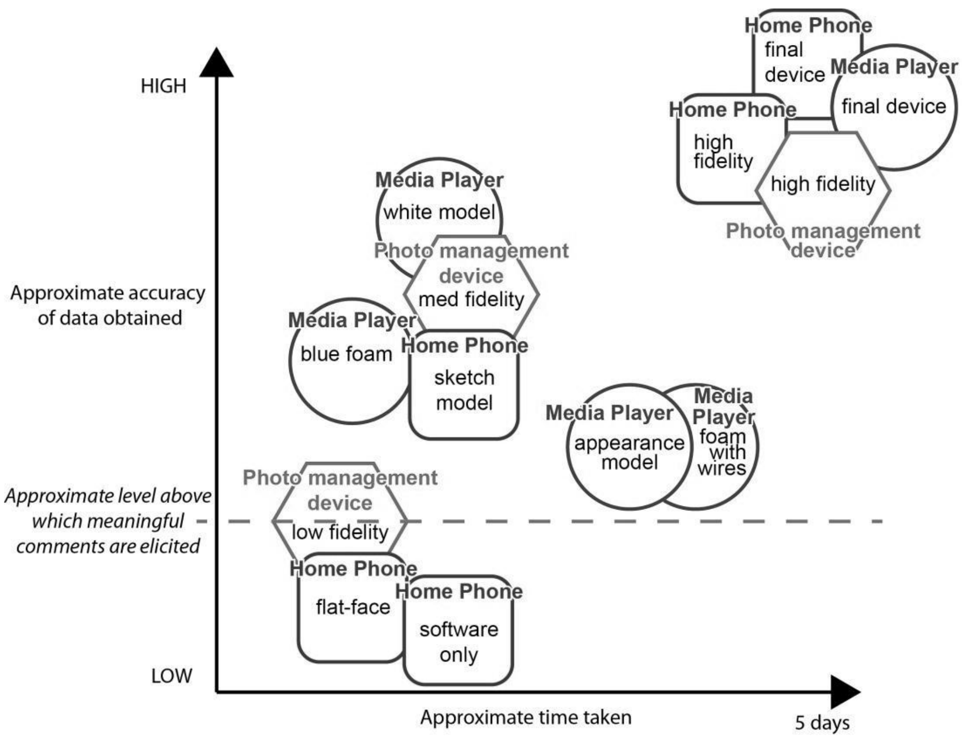


Figure 51: Relative success of the prototype versus the time taken to create the prototype

In total, eleven prototypes of three separate products were studied. The relative success of each prototype can be determined by comparing the data the prototype produced during user trials to the other prototypes in that series (including the final product if present). Figure 51 shows the relative accuracy of the data elicited by the prototype versus the time taken to create the prototype (for the Contextual Study these are approximate times). The relative accuracy of the prototype is an estimated value based on the results of each prototype within each study. In addition, an approximate line has been included above which prototypes are considered 'effective' by producing meaningful comments.

The next section will discuss the hypothesis that the least effective prototypes fall below certain levels of either active or passive physicality in relation to the design intent, and the most effective prototypes balance active and passive phsyciality equally. It will then discuss the two 'classes' of prototypes idenitifed in the literature review; prototypes without embedded electronics, and 'smart' prototypes.

6.5.1 The Lower Limits of Active and Passive Physiclaity

Figure 51 shows an approximate line below which prototypes seem to become less effective at producing meaningful comments. The least successful prototypes did not address one, or both, aspects of physicality. These are: the 'flat-face' and 'software-only' prototypes of the Contextual Study, and the 'low fidelity' prototype of Study One. Both the 'flat-face' and 'low fidelity' prototypes are 'non-automatic' or facilitator driven, while the 'software-only' prototype has no passive physicality implemented. This shows that prototypes should be 'automatic' for user studies which focus on the complete design (not just the interface). However, the 'blue foam' prototype of Study Two was facilitator driven, yet its implementation enabled users to feel like they were interacting in an 'automatic' way. This was because the digital interface was triggered by the facilitator rather than the facilitator manually updating individual screens.

6.5.2 Balancing Physicality

In the studies, the most successful prototypes balanced both active and passive physicality equally, these included the very low fidelity 'sketch' prototype and 'high-fidelity' prototype of the Contextual Study, and the low fidelity 'blue foam' prototype and the higher fidelity 'white model' of Study Two. The 'appearance' and 'blue foam with wires' prototypes of Study Two focused too much on either active or passive physicality with less consideration of the other, although these prototypes produced usable data, it was not as reliable as the well-balanced prototypes.

When the prototypes are considered in terms of physicality and compared to the time taken to create each prototype it can be seen that the extra time invested in some of the prototypes was perhaps inefficient. Both the 'foam model with wires' and 'appearance' models of Study Two took more time than the 'white model' of that same study but were less successful. In order to increase the effectiveness of the 'appearance' prototype and 'foam model with wires', they could be combined with further investment to source buttons and dials that did not have a significant impact on passive physicality. This type of investment would be more justifiable towards the later stages of the design process.

As was hypothesized in the Contextual Study, it is not the level of fidelity that is important in these prototypes. Rather it is considerations about the physicality of the prototype in relation to the design intent, specifically that there is a good balance between active and passive elements of the prototype as identified by Study One and Study Two.

6.5.3 Prototypes without embedded electronics

The 'software-only' prototype of the Contextual Study had no physicality, although it could be argued that there is some physicality if the intended product is touchscreen and it is prototyped on a touchscreen. However, this was not the case in this study and the inclusion of an 'appearance' model can be used to address this. The 'appearance' prototype (Study Two) was close to the design intent for passive physicality but active physicality remained low. Previously, it was thought that the inclusion of this 'appearance' model was adequate to inform the design process but these studies have demonstrated that these prototypes produced unreliable data compared to the other prototypes in each study. This is supported by Blackler (2009) who noted that when the prototype was operated on a touchscreen this seemed to encourage participants to think that the product would also work by touchscreen, even for buttons that were clearly depicted as physical. Indeed Ehn & Kyng (1991, p. 193) warn that in computer-supported prototypes; "*The closer the two 'roles' get, and the less familiar the computer is, the more careful one had to be in avoiding attributing the wrong aspects of the mock-up … to the future product.*".

In Study Two, the interface required between 59% and 96% of prototyping time depending on its level of fidelity. This is supported by Woolley (2008), who found that greater length of time was spent prototyping the software interface than creating the physical prototypes. Therefore, with as little as 4% extra time (in the case of the Media Player of Study Two), the effectiveness of the prototype can be greatly improved by the inclusion of an interactive physical model bringing the level of active versus passive physicality into balance.

The literature review identified that a majority of the research on fidelity concentrated on software only products (Section 2.4.3). The research suggests models of fidelity (such as the 'five dimensions of fidelity' (McCurdy, Connors, Pyrzak, Kanefsky, & Vera, 2006)) can be used to ensure the appropriate aspects of the software are prototyped. Indeed much of this research can be directly applied to computer embedded products. The Home Phone Contextual Study had the greatest difference in the functionality of the interface, the 'high fidelity' and 'software-only' prototypes had considerably more time spent on the functionality of the interface compared to the 'sketch' and 'flat-face' prototypes. The results show that considerations of physicality had more effect on users' feedback than the fidelity of the interface. This indicates that the implementation of the interface should be considered in terms of both time and fidelity.

The effectiveness (or lack thereof) of paper prototyping has been a surprise outcome of these studies. It is a technique used regularly in commercial work, yet the 'low-fidelity' prototype of Study One shows that the lack of real-time feedback (active physicality) results in a decrease in the quality of results. This prototyping technique is classified as 'non-automatic' by Nilsson and Siponen (2006) since the facilitator plays a very noticeable role in the eyes of the participant. Sefelin *et al.* (2003) proposed that the reduction in the quality of data was due to the delay of the facilitator in updating the interface. This causes the participant to feel they are adding 'unnecessary' work for the facilitator and restricts the participants ability to 'explore' the interface. Indeed, Nielsen (1990) found a similar result where users found significantly less 'global' problems when using a paper prototype compared to a software prototype. Yet paper prototyping has been proven to be a successful method for usability studies (Snyder C. , 2003) (Sefelin, Tscheligi, & Giller, 2003). The user trials of Study One were designed to obtain feedback about the scope of the overall design (what Nielsen describes as 'global' considerations) rather than task structure. This suggests that paper prototyping could be more appropriate when exploring the detailed information architecture of an interface as opposed to global design considerations. This shows a lower limit of active physicality for early stage user feedback based on usability trials; the prototype should appear to be 'automatic' or real-time.

Study Two addressed the lack of active physicality in the paper prototype through the 'blue foam' prototype; this increased levels of active physicality through the facilitator operating an 'automatic' interface thus balancing the levels of active and passive physicality. This 'blue foam' prototype proved successful in eliciting reliable user feedback. The higher fidelity 'white model' outperformed this foam prototype but as a quick and dirty prototype this 'blue foam' prototype was a success.

The 'appearance' prototype of Study Two posed an interesting question in relation to the buttons. On this prototype the buttons felt similar to the final product but they were not functional. Active physicality has been proposed to be the perceptible feedback of interacting with the product; and, haptically at least, the interactions are accurate. Yet these interactions do not trigger any other feedback, therefore the user is not able to relate their interactions to the product as a whole and this has implications for the creation of accurate mental models.

6.5.4 'Smart' prototypes

The inclusion of electronics within the prototype is a common way to increase the fidelity of interactive prototypes; this enables real-time interaction and an improved richness of interactivity (impacting active physicality). Seven of the prototypes covered a variety of approaches to making the prototype 'smarter'. Some of those approaches have resulted in an adjustment to the physical form and some have resulted in additional wires being present; in all of the prototypes studied, the screen was outside the physical model. The Contextual Study and Study Two demonstrate that the remote screen had no impact on the data gathered by comparing results to the final product which had an integrated screen. This allows a significant reduction of the development time of prototypes pushing levels of fidelity and physicality even lower.

The prototypes that had a significant impact on passive physicality produced the least reliable data; the two extreme cases in the studies were the 'flat-face' prototype of the Contextual Study and the 'foam model with wires' of Study Two. The physical form of the 'foam model with wires' was distorted due to the size of the switches and dial used, in addition, the wires and prototyping board were clearly visible impacting passive physicality. Participants commented that they felt 'intimidated' by the appearance of the electronics and that interactions were not easy to reach on the prototype. The 'flat-face' prototype used in the Contextual Study had a paper cut-out covering the buttons to avoid the need to embed the buttons in the front of the model, saving a few hours of work. The effect of this paper cut-out was two-fold; firstly the level of passive physicality was too low because the operation of the prototype is not apparent to the user, and secondly, the 'hit-area' shown on the paper cut out was not the true hit area of the buttons beneath it (active physicality).

6.6 Conclusions

This research began with a broad investigation into the meaning of physicality in relation to computer embedded products. Physicality was determined to be the physical aspects or qualities of both an object and an interaction; this includes our physical bodies in relation to that object.

Physicality is central to our experience of computer embedded products, from how we exist in our bodies within the physical world and how we perceive interactions with the physical world to the point at which we interact with that physical world.

Three interrelated themes were identified; the meaning of physicality in relation to computer embedded products, the design process of computer embedded products, and the physical manifestations of the design process (specifically prototypes). These were explored in the literature review specifically in relation to the design of computer embedded products. In addition, three of the studies that were undertaken as part of the concurrent EPSRC/AHRC funded project (the DEPtH project) provided contextual information. The research question was informed by both the literature review and the Contextual Studies.

Research Question: Can a better understanding of physicality help in the creation of more effective low-fidelity physical interactive prototypes?

The literature review identified that of the research that does focus on the fidelity of physical interactive prototypes; the construction of the physical prototype is rarely typical of the product design process. The Contextual Study of the Mobile Home Phone proposed that it is not the level of fidelity that is important but rather the considerations of tangibility and physicality.

Two studies were undertaken to explore a range of prototypes of two different products; a conceptual photo management product and a media player. Through these studies, the notion of active and passive physicality was proposed where; passive physicality is the **perceived affordance based on the tangibility and visual appearance of the prototype,** and active physicality is the **perceptible experience of interacting with the prototype.**

A hypothesis was developed whereby both active and passive physicality can be considered on a scale of low to high and that prototypes which fall below certain levels of either active or passive physicality in relation to the design intent are least effective and prototypes that balance active and passive physicality equally are the most effective. In this situation, an 'effective' prototype is one which elicits feedback related to the intended design to enable the next iteration of the design to take place.

The proposal that active and passive physicality should be 'balanced' recognises that many prototyping construction techniques require a compromise of some kind. For example, the use of electronics within a prototype necessitates components and power requirements which could impact the size of the prototype and the demand for a highly realistic prototype could impact the way in which the prototype can be interacted with. In these scenarios the resultant physicality of the prototype is affected even though its fidelity is not necessarily altered. Without an understanding of this affect any prototype created will be limited in its effectiveness.

The findings of the Contextual Study and both user studies were then re-examined. The Contextual Study demonstrates how an understanding of active and passive physicality can provide a framework by which to better understand unforeseen results. Study One (the conceptual photo management product) demonstrates that physicality is not solely dependent on fidelity. Study Two (the media player) provides an insightful example of the boundaries of active and passive physicality.

When the prototypes from all the studies are compared in relation to the effectiveness of each prototype as informed by the user trials, it can be seen that the most effective prototypes do balance active and passive physicality and that the least effective prototypes fall below a certain limit of either active or passive physicality. These limits are the level of automation for active physicality, and the inclusion of a physical prototype indicative of the form of the intended design for passive physicality.

6.6.1 Future Work

Future work will seek input about the relevance of passive and active physicality from industry and other academics. In addition to this, prototypes emerging from both research and commercial projects can be evaluated against the notions of passive and active physicality to gain general feedback about the application of the notions and also specific feedback about the relevance of notions of physicality when there are no prototypes to make comparisons against.

The prototypes presented in this thesis are all of handheld computer embedded products. Passive physicality has been defined as being the perceived affordance based on the tangibility and visual appearance of the prototype, yet if that product was not intended to be handheld does that description still hold true? The user still physically interacts with the device, therefore the location of interactions is still important but the way in which the device is held is not. Future work in this area could apply these notions of passive and active physicality to computer embedded products in general.

The prototypes used in this thesis are button and dial based. Further case studies could focus on different interaction technologies such as gesture based interactions and products that change shape such as those with flexible screens.

Many more prototyping techniques exist than have been utilised in this research. Future work could explore prototyping techniques such as augmented reality and virtual prototyping.

In this thesis the 'effectiveness' of the prototype was determined by its performance in a typical early stage user trial. A prototype might be constructed for a number of reasons, user testing being just one of many. Therefore, if a prototype is intended to be used for group discussion, for example, do the notions of active and physicality still hold true? In this scenario the facilitator is able to supplement the prototype with verbal description and additional props. Future work in this area could apply notions of passive and active physicality to a wider range of user focused research.

6.6.2 Contribution to Knowledge

The contribution to knowledge of this work is the framework of active and passive physicality which can be used to construct more effective interactive prototypes for user trials.

The framework demonstrates that both active and passive physicality should be considered on a scale of low to high and that prototypes which fall below certain levels of either active or passive physicality in relation to the design intent are least effective and prototypes that balance active and passive physicality equally are the most effective.

This framework is shown graphically in Figure 52. The two axes show the extent of active and passive phsycialty with the orgin being 'no physiclaity' and the 'highest' physiclaity being equivalent to the final design intent. The areas below the dashed lines show that there it a point at which physiclaity can be considered too low for an effective prototype. The diagonal depicts where prototype considered to have 'balanced physicilaty' should fall on the graph.

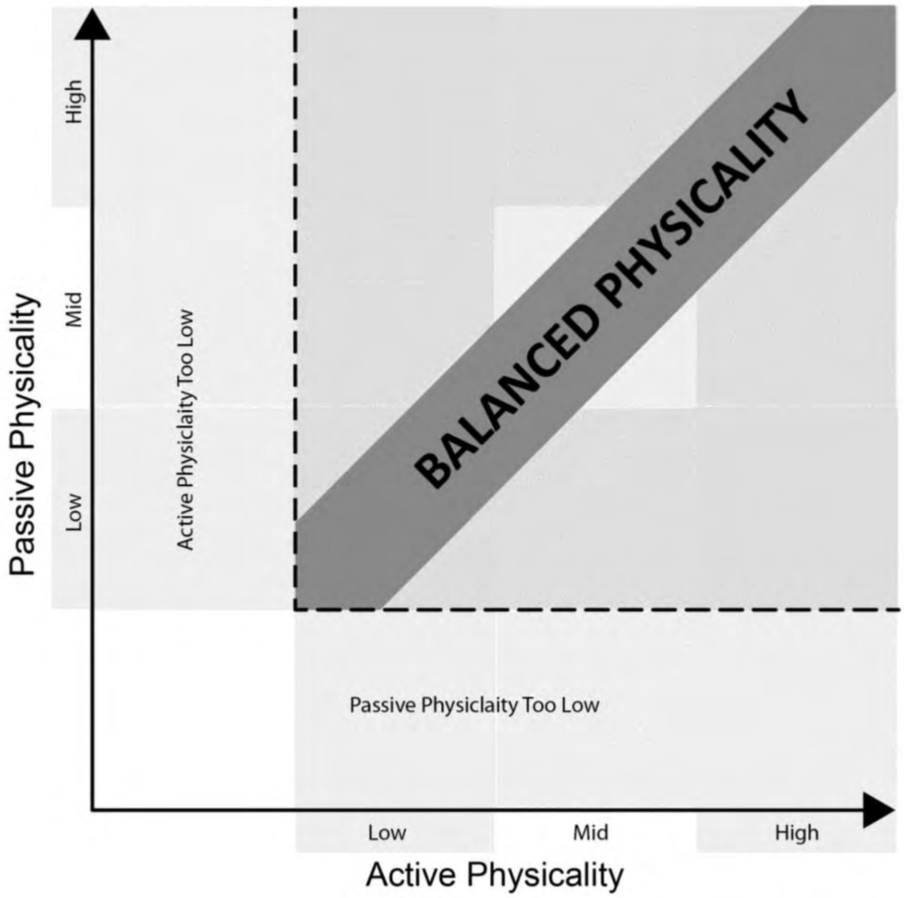


Figure 52: Framework of Active and Passive Physicality

6.6.3 Key Findings

In addition to the development of the framework of active and passive physicality, key findings have also emerged from this research:

1. Physicality is not solely dependent on fidelity
2. Paper prototyping is not as effective in user trials which focus on the complete design (not just the interface)
3. Screen-only prototypes are not effective in this type of user trial
4. Use of electronic prototyping toolkits should include physical model making
5. Creating facsimile 'appearance models' requires additional electronic interaction.

Glossary

Affordances

The term affordance is discussed in detail in Section 2.2.2 starting on page 11. The use adopted in this thesis is that of Norman's perceived affordances. Perceived affordances are the qualities of an artefact that suggest the possibility of interaction in relation to the capabilities to the user, these are affected by the perceptual and mental capabilities of that person.

Fidelity

Fidelity is discussed in detail in Section 2.4.3 starting on page 53. The term fidelity is used to convey how realistic an artefact is in relation to the final design intent. The scale 'low' to 'high' is typically used when describing fidelity with a high fidelity artefact being very close to the design intent and a low fidelity prototype only demonstrating a limited proportion of the design intent.

Physicality

Physicality is the physical aspects or qualities of both an object and interaction; this includes our physical bodies in relation to that object.

For this thesis physicality, like fidelity, is considered on a scale of 'low' to 'high' where an artefact with high physicality is very close to the final design intent and a low physicality artefact only represents the design intent in a limited manner.

Active Physicality

"The perceptible experience of interacting with the prototype". Active physicality results from the user doing something to the prototype that requires a system response or mechanical action such as pressing a button or performing a gesture.

Passive Physicality

"The perceived affordance of a prototype mainly based on the visual appearance and tangibility of the prototype". Passive physicality can be determined predominantly by looking at and picking up a prototype, if there are any other features of the intended design such as lights, smells, textures then these should be considered as passive physicality unless the user has been required to initiate an action.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-Centered Design. In W. Bainbridge, *Encyclopedia of Human-Computer Interaction.* Thousand Oaks: Sage Publications.

Bartneck, C., & Rauterberg, M. (2007). HCI reality - an unreal tournament? *International Journal on Human-Computer Studies, 65*(8), pp. 737-743.

Baxter, M. (1995). *Product design: a practical guide to systematic methods of new product development.* Florida, USA: CRC Press.

Bdeir, A. (2009). Electronics as material: littleBits. *Proceedings of the Third International Conference on Tangible and Embedded Interaction (TEI'09)* (pp. 397-400). Cambridge, UK: ACM.

Bergman, E., & Haitani, R. (2000). Designing the PalmPilot: A Conversation with Rob Haitani. In E. Bergman, *Information appliances and beyond: interaction design for consumer products* (pp. 81-102). Morgan Kaufmann.

Blackler, A. (2009). Applications of high and low fidelity prototypes in researching intuitive interaction. *Undisciplined! Design Research Society Conference 2008.* Sheffield.

Bordegoni, M., & Cugini, U. (2005). Design Products with your hands. *In Proceedings of Virtual Concept*, *Vol 1.* Biarritz, France.

Bowles, T. (2011, February 25). *Apple, UCD, and Innovation*. Retrieved October 13, 2013, from the Human Factors Blog: http://humanfactorsblog.org/2011/02/25/apple-ucd-and-innovation-a-guest-post-by-travis-bowles

Boztepe, S. (2007). User Value: Competing Theories and Models. *International Journal of Design, 1*(2), 55-63.

Branham, R. (2000). Given the radically changing work environment and new worldviews, what kinds of new 'tools' do designers need to survive and successfully deal with tomorrow's design problems. *Proceedings of IDSA 2000 National Education Conference*, (pp. 18-20). Louisiana, USA.

Broek, J. J., Horváth, I., & Lennings, A. F. (2000). Using Physical Models in Design. *CAID 2000*, (pp. 155-163).

Brunel University London. (2014, June 12). *Industrial Design and Technology BA*. Retrieved June 16, 2014, from Brunel University London: http://www.brunel.ac.uk/courses/undergraduate/industrial-design-and-technology-ba

Burleson, W., Jensen, C., Raaschou, T., & Frohold, S. (2007). Sprock-it: a physically interactive play system. *Proceedings of the 6th international conference on Interaction design.* Aalborg, Denmark: ACM.

Busher, M. (2006). Making the Digital Palpable. *The proceedings of The First International Workshop on Physicality*, (pp. 1-2). Lancaster.

Buxton, B. (2007). *Sketching User Experiences: Getting the Design Right and the Right Design.* London: Morgan Kaufmann.

Carroll, J. M. (2013). Human Computer Interaction - brief intro. In M. a. Soegaard, *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Aarhus, Denmark: The Interaction Design Foundation.

Clark, A. (1998). Where Brain, Body and World Collide. *Daedalus: Journal of the American Academy of Arts and Science (Special Issue on the Brain), 127*(2), p.257-280.

Cockton, G. (2013). Usability Evaluation. In M. a. Soegaard, *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Aarhus, Denmark: The Interaction Design Foundation.

Craik, K. J. (1967). *The Nature of Explanation.* Cambridge University Press Archive.

Culverhouse, I. (2011). *Investigation into the insights generated through the application of interactive prototyping during the early stages of the design process (PhD thesis).* Cardiff: UWIC.

Debian. (n.d.). Retrieved 10 30, 2013, from Debian: http://www.debian.org/

Dieter, G. E. (2000). *Engineering Design: a materials and processing approach (3rd edition).* USA: McGraw-Hill.

Digital Trends. (2009). *Digital Trends: CES 2009 iRiver Spinn.* Retrieved 2009, from http://www.youtube.com/watch?v=gNmichwkWww&feature=related

Dix, A., Gill, S., Ramduny-Ellis, D., & Hare, J. (2010). Design and Physicality – Towards an Understanding of Physicality in Design and Use. In T. Inns, *Designing for the 21st Century: Interdisciplinary Methods and Findings* (pp. 172-189). Gower Publishing, Ltd.

Djajadiningrat, T., Overbeeke, K., & Wensveen, S. (2002). But how, Donald, tell us how? *Designing Information Systems (DIS)*, (pp. 285-291). London.

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction.* Cambridge, MA.: MIT Press.

Dumas, J. S., & Redish, J. (1999). *A Practical Guide to Usability Testing.* Intellect Books.

Ehn, P. (1988). *Work Oriented Design of Computer Artifacts.* Hillsdale, NJ: Lawrence Erlbaum.

Ehn, P., & Kyng, M. (1991). Cardboard Computers: Mocking-it-up or Hands-on the Future. In J. Greenbaum, & M. Kyng, *Design at work: Cooperative design of computer systems* (pp. 169-195).

Evans, M. A. (2005). Rapid prototyping and industrial design practice: can haptic feedback modelling provide the missing tactile link? *Rapid Prototyping Journal*, 153–159.

Floyd, C. (1984). A Systematic Look at Prototyping. In R. Budde, *Approaches to prototyping* (pp. pp 1-18). Springer.

Frens, J. W. (2006). Designing for rich interaction: Integrating form, interaction and function.

Gaver, W. (1991). Technology Affordances. *SIGCHI conference on Human factors in computing systems Reaching through technology*, (pp. p. 79-84).

Gaver, W., Beaver, J., & Benford, S. (2003). Ambiguity as a resource for design. *Proceedings of the ACM-SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*, (pp. 233-240). Fort Lauderdale, FL.

Geomagic. (n.d.). *Geomagic Touch X*. Retrieved 05 14, 2014, from Geomagic: http://www.geomagic.com/en/products/phantom-desktop/overview/

Ghazali, M. (2006). Physicality and something else it's associated with. *The First International Workshop on Physicality*, (pp. 35-36).

Ghazali, M., & Dix, A. (2005). Knowledge of Today for the Design of Tomorrow. *Proceedings of the 2nd International Design and Engagibility Conference (IDEC).* Edinburgh, UK.

Ghazali, M., Ramduny-Ellis, D., Hornecker, E., & Dix, A. (2006). Proceedings of the First International Workshop on Physicality. *Physicality 2006.* Lancaster University, UK.

Gibson, J. (1977). *"The theory of affordances" in Perceiving, Acting, and Knowing.* Hoboken, NJ: John Wiley & Sons Inc.

Gibson, J. (1986). *The Ecological Approach to Visual Perception.* Psychology Press; New edition.

Gill, S. (2013). Computer-embedded Design: PAIPR Prototyping. In L. Valentine (Ed.), *Prototype: Craft in the Future Tense.* Bloomsbury.

Gill, S., Loudon, G., Woolley, A., Hare, J., Walker, D., Dix, A., et al. (2008). Rapid development of tangible interactive appliances: achieving the fidelity/time balance. *International Journal of Arts and Technology, 3*(3-4), 309-331.

Goltz, S. (2012, May 10). *UXD / IXD App Comparison (Prototyping and Wireframing).* Retrieved May 22, 2013, from https://docs.google.com/spreadsheet/ccc?key=0AlS4P_cuEIKDdFpDVEpaalBDQzJjclpzdGFDNVJvaUE#gid=0

Goodwin, K. (2009). *Designing for the Digital Age: How to Create Human-Centered Products and Services.* John Wiley & Sons.

Gould, J. D., & Lewis, C. (1983). Designing for Usability -Key Principles and what Designers Think. *CHI '83* (pp. 50-53). Boston: ACM Press.

Gould, J. D., Conti, J., & Hovanyecz, T. (1982). Composing Letters with a Simulated Typewriter. *CHI '82 Proceedings of the 1982 Conference on Human Factors in Computing Systems* (pp. 367-370). ACM.

Greenberg, S., & Fitchett, C. (2001). Phidgets: easy development of physical interfaces through physical widgets. *Proceedings of the 14th annual ACM symposium on User interface software and technology.* ACM, Orlando, Florida.

Gribnau, M. (1999). *Two-handed interaction in computer supported 3D conceptual modelling.* Doctoral Dissertation, Delft University of Technology.

Hall, R. R. (1999). Usability and Product Design: a Case Study. In W. S. Green, & P. W. Jordan, *Human Factors in Product Design: Current Practice and Furture Trends* (pp. 85 - 91). Tylor & Francis.

Harrelson, D. (2009, September 16). *Rapid Prototyping Tools Revisited.* Retrieved October 30, 2013, from Adaptive Path: https://spreadsheets.google.com/pub?key=pOa2Uqiakxlry5hNuZm89Eg&output=html

Hart, C. (1943). The Hawthorne Experiments. *Canadian Journal of Economics and Political, 9*(2), 150–63.

Hartmann, B., Klemmer, S. R., Bernstein, M., & Mehta, N. (2005). d.tools: Visually Prototyping Physical UIs through Statecharts. *in Extended Abstracts of UIST 2005.*

Hartson, H. R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology, 22*(5), 315-338.

Haugeland, J. (1998). Mind Embodied and Embedded. In J. Haugeland, *Having Thought.* Cambridge, MA: Harvard University Press.

Heskett, J. (1980). *Industrial Design.* London: Thames & Hudson.

Holtzblatt, K., Burns Wendell, J., & Wood, S. (2005). *Rapid Contextual Design.* Elsevier Inc.

Israel, J. H. (2007). Physical Interaction: The Basis of Human-Computer-Interaction. *Proceedings of the Second International Workshop on Physicality* (pp. 83-86). Cardiff: UWIC Press.

Jobs, S. (1998, May 12). Steve Jobs on Apples' resurgence: "Not a one-man show". (BusinessWeek, Interviewer)

Jørgensen, M. W., & Phillips, L. J. (2002). *Discourse Analysis as Theory and Method.* SAGE.

Kelley, T. (2001). *The Art of Innovation: Lessons in Creativity from IDEO, America's Leading Design Firm.* Broadway Business.

Kelly, M. (1998). *Encyclopedia of Aesthetics.* Oxford University Press.

Kelly, M. (2007, August). *Interactive Prototypes with PowerPoint.* Retrieved 10 29, 2013, from boxesandarrows: http://boxesandarrows.com/interactive-prototypes-with-powerpoint

Kim, J., & Wilemon, D. (2002). Focusing the fuzzy front-end in new product development. *R&D management, 32*(4), 269-279.

Kim, J.-W., & Nam, T.-J. (2013). EventHurdle: supporting designers' exploratory interaction prototyping with gesture-based sensors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)* (pp. 267-276). New York, NY, USA: ACM.

Kjeldskov, J., & Graham, K. (2003). A Review of Mobile HCI Research Methods. *Proceedings of the 5th International Symposium on Human-Computer Interaction with Mobile Devices and Services (Mobile HCI 2003)*, (pp. 317-355).

Kuniavsky, M. (2003). *Observing the User Experience: A Practitioner's Guide to User Research.* San Francisco, CA: Morgan Kaufmann.

Larssen, A. T., Robertson, T., & Edwards, J. (2006). How it Feels, not Just How it Looks: When Bodies Interact with Technology. *OZCHI 2006.* Sydney, Australia.

Lennings, A. F., Broek, J. J., Horváth, I., Sleijffers, W., & Smit, B. d. (2000). Editable Physical Models for Conceptual Design. *International Symposium series on Tools and Methods of Competitive Engineering (TMCE) 2000* (pp. 665-674). Delft University Press.

Lewis, C., & Rieman, J. (1993). *Task-centered User Interface Design: A Practical Introduction.* University of Colorado, Boulder: This shareware book is available at ftp. cs. colorado. edu.

Lim, Y.-k., Pangam, A., Periyasam, S., & Aneja, S. (2006). Comparative Analysis of High- and Low-fidelity Prototypes for More Valid Usability Evaluations of Mobile Devices. *NordiCHI 2006.*

Lim, Y.-K., Stolterman, E., & Tenenberg, J. (2008, July). The Anatomy of Prototypes: Prototypes as Filters, Prototypes as Manifestations of Design Ideas. *ACM Transactions on Computer-Human Interaction, 15*(2), 7:0 - 7:27.

Liu, L., & Khooshabeh, P. (2003). Paper or Interactive? A Study of Prototyping Techniques for Ubiquitous Computing Environments. *CHI 2003: NEW HORIZONS* (pp. 1030-1031). Ft. Lauderdale, Florida, USA.: ACM.

Lowgren, J. (2013). Interaction Design - brief intro. In M. a. Soegaard, *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Aarhus, Denmark: The Interaction Design Foundation.

McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B., & Vera, A. (2006). Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success. *CHI '06 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1233-1242). ACM.

McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and Evolving a Concept. *Graphics Interface 2000.* Montreal.

Merleau-Ponty, M. (1945). *Phenomenology in Perception.* (D. A. Landes, Trans.) Paris.

Merriam-Webster. (n.d. (a)). *Physical*. Retrieved 09 16, 2013, from Merriam-Webster.com: http://www.merriam-webster.com/dictionary/physical

Merriam-Webster. (n.d. (b)). *Prototype*. Retrieved October 29, 2013, from Merriam-Webster.com: http://www.merriam-webster.com/dictionary/prototype

Miller, P. (2009). *iriver SPINN review*. Retrieved 2009, from Engadget: http://www.engadget.com/2008/10/16/iriver-spinn-review/

Moggridge, B. (2007). *Designing Interactions.* Cambridge, MA.: The MIT Press.

Molich, R., & Dumas, J. S. (2008). Comparative usability evaluation (CUE-4). *Behaviour & Information Technology, 27*(3), pages 263-281.

Molich, R., Bevan, N., Butler, S., Curson, I., Kindlund, E., Kirakowski, J., et al. (1998). Comparative evaluation of usability tests. *Usability Professionals Association Conference.* Washington, DC.

Molich, R., Jefferies, R., & Dumas, J. .. (2007). Making Usability Recommendations Useful and Usable. *Journal of Usability Studies*, 162-179.

Narayan, K. L., Rao, K. M., & Sarcar, M. (2008). *Computer Aided Design and Manufacturing.* PHI Learning Pvt. Ltd.

Neilsen, J. (1990). Paper versus computer implementation as mockup scenarios for heuristic evaluation. *Interact '90*, (pp. 315-320). Cambridge, UK.

Nielsen, J. (1993). *Usability Engineering.* San Francisco, CA.: Academic Press.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings of ACM INTERCHI'93 Conference* (pp. 206-213). Amsterdam, The Netherlands: ACM.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Proceedings of CHI '90: Conference of Human Factors in Computing Systems*, (pp. 249-256). Seattle, WA.

Nilsson, J., & Siponen, J. (2006). Challenging the HCI Concept of Fidelity by Positioning Ozlab Prototypes. In *Advances in Information Systems Development* (pp. 349-360). Springer US.

Nokia. (2011, 03 03). *Protoyping your application through Flowella*. Retrieved 10 30, 2013, from Nokia Developer: http://developer.nokia.com/community/wiki/Archived:Prototyping_your_application_through_Flowella

Noldus. (n.d.). *Observer-XT*. Retrieved 07 15, 2013, from Noldus: http://www.noldus.com/human-behavior-research/products/the-observer-xt

Norman, D. (1983). Some observations on mental models. In D. Gentner, & A. L. (eds), *Menatl Models.* Lawrence Earlbaum Associates.

Norman, D. (1988). *The Psychology of Everyday Things.* Basic Books.

Norman, D. (1998). *The Design of Everyday Things.* New York: Basic Books.

Norman, D. (1999). *The Invisible Computer: Why good products can fail, the personal computer is so complex, and information appliances are the solution.* Cambridge, MA: MIT Press.

Norman, D. (2004). *Affordances and Design.* Retrieved 10 3, 13, from JND: http://www.jnd.org/dn.mss/affordances_and.html

Norman, D. (2013). Commentary on: Svanaes, Dag (2013): Philosophy of Interaction. *"The Encyclopedia of Human-Computer Interaction, 2nd Ed."*. (M. a. Soegaard, Ed.) Aarhus, Denmark.

Overbeeke, K., & Hummels, C. (2013). Industrial Design. In M. a. Soegaard, *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Aarhus, Denmark: The Interaction Design Foundation.

Ozenc, F. K., Kim, M., Zimmerman, J., Oney, S., & Myers, B. A. (2010). How to support designers in getting hold of the immaterial material of software. *CHI '10 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. Pages 2513-2522).

Pering, C. (2002, November & December). Interaction design prototyping of Communicator devices: Towards meeting the hardware-software challenge. *Interactions, 9*(6), 36-46.

Pevsner, N. (2005). *Pioneers of Modern Design: From William Morris to Walter Gropius.* Yale University Press.

Potter, N. (1969). *What Is a Designer: Things, Places, Messages.* Studio Vista.

Pugh, S. (1991). *Total design: integrated methods for successful product engineering.* Wokingham: Addison-Wesley.

Raizman, D. (2010). *A History of Modern Design: Graphics and Products Since the Industrial Revolution.* London : Laurence King.

Ramduny-Ellis, D., Dix, A., Hare, J., & Gill, S. (2007). Proceedings of the Second International Workshop on Physicality. *Physicality 2007.* Lancaster University: UWIC Press.

Ramduny-Ellis, D., Dix, A., Hare, J., & Gill, S. (2009). Physicality 2009 - towards a less-GUI interface . *Physicality 2009.* Cambridge.

Redish, J., Bias, R. G., Bailey, R., Molich, R., Dumas, J., & Spool, J. M. (2002). Usability in practice: formative usability evaluations - evolution and revolution. *CHI '02 Extended Abstracts on Human Factors in Computing Systems* (pp. Pages 885-890). New York: ACM.

Reeves, S. (2006). Physicality, spatial configuration and computational objects. *The First International Workshop on Physicality .*

Reinertson, D. G. (1999). Taking the Fuzziness out of the Fuzzy Front End. *Research Technology Management*, pp 25-31.

Rogers, E. M. (2010). *Diffusion of Innovations* (Fourth ed.). New York: Simon & Schuster Inc.

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction Design: Beyond Human-Computer Interaction* (Third ed.). John Wiley & Sons.

Römer, A., Pache, M., Weißhahn, G., Lindemann, U., & Hacker, W. (2001). Effort-saving product representations in design -results of a questionnaire survey. *Design Studies, 22*(6), 473-491.

Rooden, M. J. (1999). Prototypes on Trial. In W. S. Green, & P. W. Jordan, *Human Factors in Product Design: Current Practice and Furture Trends* (pp. 138 - 150). Tylor & Francis.

Rosson, M. B., & Carroll, J. M. (2002). *Usability Engineering: Scenario-based Development of Human-computer Interaction.* San Francisco: Morgan Kaufmann.

Rosson, M. B., Ballin, J., & Rode, J. (2005). Who, what, and how: a survey of informal and professional Web developers. *IEEE Symposium on Visual Languages and Human-Centric Computing*, (pp. 199 - 206).

Rubin, J. (1994). *Handbook of Usability Testing.* Indianapolis, Indiana: John Wiley & Sons.

Rudd, J., Stern, K., & Isensee, S. (1996, January). Low vs. high-fidelity prototyping debate. *Interactions, 3*(1), 76-85.

Sade, S., Nieminen, M., & Riihiaho, S. (1998). Testing usability with 3D paper prototypes-Case Halton system. *Applied Ergonomics, 29*(1), 61-13.

Saffer, D. (2010). *Design for interaction: Creating innovative applications and devices.* Berkeley, CA.: New Riders.

Sanders, M. S., & McCormick, E. J. (1992). *Human Factors in Engineering and Design* (Seventh ed.). McGraw-Hill Inc.

Sauer, J., & Sonderegger, A. (2009). The influence of prototype fidelity and aesthetics of design in usability tests: Effects on user behaviour, subjective evaluation and emotion. *Applied Ergonomics, 40*, 670–677.

Sauer, J., Franke, H., & Ruettinger, B. (2008). Designing interactive consumer products: Utility of paper prototypes and effectiveness of enhanced control labelling. *Applied Ergonomics*, 71–85.

Sauer, J., Seibel, K., & Ruttinger, B. (2010). The influence of user expertise and prototype fidelity in usability tests. *Applied Ergonomics, 41*(1), 130–140.

Scali, S., Shillito, A. M., & Wright, M. (2002). Thinking in space: concept physical models and the call for new digital tools. *Crafts in the 20th Century.* Edinburgh .

Schrage, M. (1999). *Serious Play: How the worlds best companies simulate to innovate.* Cambridge, MA: Harvard Business School Press.

Sefelin, R., Tscheligi, M., & Giller, V. (2003). Paper prototyping - what is it good for?: a comparison of paper- and computer-based low-fidelity prototyping. *CHI '03: New Horizons, Short Talk* (pp. 778-779). ACM.

Shillito, A. M. (2004). The TACITUS experience: a new spatial, multi-sensory digital interface supporting creativity. *PixelRaiders2.* Sheffield.

Silver, J., & Rosenbaum, E. (2012). Makey Makey: Improvising Tangible and Nature-Based User Interfaces. *proceedings of the International Conference on Tangible, Embedded and Embodied Interaction.*

Snyder, C. (2003). *Paper Prototyping,: The Fast and Easy Way to Design and Refine User Interfaces.* San Francisco, CA.: Morgan Kaufmann.

Snyder, C. (2006). Bias in usability testing. *Boston Mini-UPA Conference.*

Spool, J., & Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. *CHI 2001.*

Stacey, M., Eckert, C., & McFadzean, J. (1999). Sketch Intrepretation in Design Communication. *International Conference on Enginering Design, ICED.* Munich.

Suwa, M., & Tversky, B. (2002). External representations contribute to the dynamic construction of ideas. In M. Hegarty, & B. M. Narayanan, *Diagrammatic Representation and Inference* (pp. 341-343). Berlin / Heidelberg: Springer.

Svanaes, D. (2013). Philosophy of Interaction. In M. a. Soegaard, *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Aarhus, Denmark: The Interaction Design Foundation.

Temkin, B. D. (2010). *Mapping The Customer Journey.* Forrester Research.

Thimbleby, H. (1991). Can anyone work the video? *New Scientist*, pp. 40-43.

Transana. (n.d.). Retrieved 07 15, 2013, from http://www.transana.org/

Treadaway, C. (2007). Translating Experience. *The Second International Workshop on Physicality* (pp. 1-6). Cardiff: UWIC Press.

Trudeau, N. (1995). *Professional modelmaking - A handbook of techniques, and materials for architects and designers.* New York: Whitney Library of Design.

Tullis, T. (1985). Designing a menu-based interface to an operating system. *Proceedings of the ACM CHI'85 Conference* (pp. 79-84). San Francisco: ACM.

Tullis, T., & Albert, B. (2008). *Measuring The User Experience: collecting, analysing and presenting usability metrics.* Morgan Kaufman.

Ulrich, K. T., & Eppinger, S. D. (2012). *Product Design and Development* (Fifth ed.). New York: McGraw-Hill International Edition.

Verlinden, J., Wiegers, T., Vogelaar, H., Horváth, I., & Vergeest, J. (2001). Exploring Conceptual Design Using a Clay-based Wizard of Oz Technique. *In 8th IFAC/IFIP/FORS/IEA symposium on analysis, design and evaluation of human-machine systems*, (pp. pp. 211-216).

Vermeeren, A. P. (1999). Designing Scenarios and Tasks for User Trials of Home Electronic Devices. In G. W. P.W, *Human Factors in Product Design: Current Practice and Future Trends* (pp. pp. 47-55). London: Taylor and Francis.

Villar, N., Lindsay, A. T., & Gellersen, H. (2005). Pin & Play & Perform: a rearrangeable interface for musical composition and performance. *NIME '05 Proceedings of the 2005 conference on New interfaces for musical expression* (pp. 188 - 191). ACM.

Virzi, R. A. (1989). What can you Learn from a Low-Fidelity Prototype? *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 224-228). Santa Monica, CA: Sage.

Virzi, R., Sokolov, J., & Karis, D. (1996). Usability problem identification using both low and high-fidelity prototypes. *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground.* Vancouver, British Columbia, Canada: ACM.

Vlaskovits, P. (2011, August 29). *Henry Ford, Innovation, and That "Faster Horse" Quote*. Retrieved October 29, 2013, from Harvard Business Review: http://blogs.hbr.org/2011/08/henry-ford-never-said-the-fast/

Von Hippel, E. (1986). Lead Users: A Source of Novel Product Concepts. *Management Science, 32*(7), 791–806.

Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes.* WW Norton.

Wixon, D., Holtzblatt, K., & Knox, S. (1990). Contextual Design: An Emergent View of System Design. *Proceedings of CHI '90: Conference of Human Factors in Computing Systems*, (pp. 329-336). Seattle, WA.

Woods, D., & Fassnacht, C. (2007). Transana v2.20. [Computer software] http://transana.org. . Madison, WI: The Board of Regents of the University of Wisconsin System.

Woolley, A. (2008). *Contextual testing of interactive prototypes at the early stages of the design process (PhD thesis).* Cardiff, Wales: University of Wales.

Wright, I. (1998). *Design Methods in Engineering and Product Design.* Maidenhead, UK: McGraw-Hill.

Chapter 11. Appendices

The appendices contain the following documents:

Appendix 1. Journal publication, Physigrams

Dix, Ghazahil, Gill, Hare, & Ramduny-Ellis (2009), **Physigrams: modelling products for natural interaction**. In Formal Aspects of Computing, Volume 21, Number 6, December, 2009.

Appendix 2. Conference publication exploring physicality in the design process

Ramduny-Ellis, Hare, Dix, & Gill (2008), **Exploring Physicality in the Design Process.** In the proceedings of the Design Research Society Conference 2008.

Appendix 3. Journal publication exploring physicality in the design process

Ramduny-Ellis, Hare, Dix, Evans, & Gill (2009), **Physicality in Design: an exploration.** In The Design Journal 13(1), pp 172-189.

Appendix 4. Journal publication exploring low-fidelity prototypes of a home phone

Gill, Walker, Loudon, Dix, Woolley, Ramduny-Ellis, & Hare (2008), **Rapid Development of Tangible Interactive Appliances: Achieving the Fidelity/Time Balance**. In Hornecker, E., Schmidt, A., and Ullmer, B. (eds) Special Issue on Tangible and Embedded Interaction, International Journal of Arts and Technology, Volume 1, No 3/4 pp 309-331.

Appendix 5. Conference publication of Study One (conceptual photo management product)

Hare, Gill, Loudon, Ramduny-Ellis & Dix (2009), **Physical fidelity: Exploring the importance of physicality on Physical-Digital conceptual prototyping**. In the proceedings of Human-Computer Interaction - Interact 2009, Uppsala, Sweden. Springer Berlin Heidelberg.

Appendix 6. Participant information sheet for Study One

Appendix 7. Conference publication of Study Two (media player)

Hare, Gill, Loudon & Lewis (2013), **The effect of physicality on low fidelity interactive prototyping for design practice**. In the proceedings of Human-Computer Interaction - Interact 2013, Cape Town, South Africa. Springer Berlin Heidelberg.

Appendix 8. Participant information sheet for Study Two

Appendix 9. IN PRESS: Journal publication providing an overview of active and passive physicality

Hare, Gill, Loudon & Lewis (2014), **Active and passive physicality: making the most of low fidelity physical interactive prototypes**. Accepted for publication in the Journal of Design Research. Inderscience.

DOI 10.1007/s00165-008-0099-y
BCS © 2008
Formal Aspects of Computing (2009) 21: 613–641

Formal Aspects
of Computing

Physigrams: modelling devices for natural interaction

Alan Dix[1], Masitah Ghazali[2], Steve Gill[3], Joanna Hare[3] and Devina Ramduny-Ellis[1]

[1] Computing Department, InfoLab21, Lancaster University, Lancaster LA9 5EZ, UK.
E-mail: alan@hcibook.com
[2] Department of Software Engineering, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia
[3] Cardiff School of Art and Design, UWIC, Cardiff, Wales, UK

Abstract. This paper explores the formal specification of the physical behaviour of devices 'unplugged' from their digital effects. By doing this we seek to better understand the nature of physical interaction and the way this can be exploited to improve the design of hybrid devices with both physical and digital features. We use modified state transition networks of the physical behaviour, which we call physiograms, and link these to parallel diagrams of the digital state. These are used to describe a number of features of physical interaction exposed by previous work and relevant properties expressed using a formal semantics of the diagrams. As well as being an analytic tool, the physigrams have been used in a case study where product designers used and adapted them as part of the design process.

Keywords: Physicality; Interaction modelling; Affordance; Natural interaction; Physical devices; Product design; Physigrams

1. Introduction

1.1. Motivation: understanding physical devices

For some years the authors have been interested in understanding what makes some interactions with physical devices seem 'natural' whilst others need to be carefully learnt. Part of this lies in the fact that interacting with ordinary objects in the physical world is natural to us; even as babies we reach out, explore our own hands, touch our mothers' faces, then play with balls and toys. Many aspects of physical interaction are informed by culture and things we have learnt about the technological world, but much is still common and would be equally natural if a person from two hundred or even two thousand years ago were suddenly transported to the present.

One aspect of our studies has been to try and unpack which properties of physical interaction are essential to make them comprehensible, and which can be relaxed [Dix03, GhD03, GhD05]. We can then use these properties to understand how to make digital interactions more natural, and thus inform several forms of design:

- pure digital interaction on a computer screen (although even this requires physical devices, such as the mouse)
- novel tangible interactions where physical objects are embued with digital properties (as found in tangible interaction, and ubiquitous computing)
- more mundane devices such as mobile phones or even electric kettles.

Correspondence and offprint requests to: A. Dix, E-mail: alan@hcibook.com

A critical method for us has been to analyse in detail commonly used artefacts such as an electric kettle or minidisk controller (already looking very dated!). We assume that when such devices appear easy to learn and use the designers have often embodied, either explicitly or implicitly, their own understanding of what makes interaction natural. So by detailed analysis of these everyday artefacts, we have gradually mined the experience, often tacit, of successful designers and documented this in terms of properties and types of device interaction, some of which we shall use later in this paper (e.g. *exposed state*, *bounce-back*, *compliant interaction*). This analysis and the early stages in developing a diagrammatic notation has been reported previously [GhD03, GhD05] and a later paper introduced a more explicitly formal specification to give a level of semantics to the diagrams and also allow formal statements of some properties [DGR07]. The current paper adds to this picture refining some of the formal expression, engaging in more detailed discussion of its implications, extending the work through a practical case study with product designers, and reflecting on use in novel device design.

1.2. This paper: focus and goals

In this paper we investigate the formal representation of the *interactive behaviour* of physical devices. The core idea is to consider devices 'unplugged'; that is entirely separate from any digital or other external functionality. When a mobile phone battery has run down, or a light switch is unscrewed from the wall; still they both have interactive physical behaviours: the phone buttons can be pressed, the light switch can be flipped with a finger, even though there is no resulting effect on the phone or light. Note the interaction here is directly with the device not with any digital or electronic behaviour. To represent these 'unplugged' devices, we will incrementally augment basic state transition networks (STNs) to enable them to represent increasingly complex physical properties identified in our previous work. We call the resulting diagrams *physigrams*.

Of course, this is not to say the digital effects are not important, indeed the advantage of specifiying the physical behaviour is that we can see the extent to which it is consonant with the digital behaviour. We use separate but connected models of physical devices (the physigrams) and of their digital effects, and so are able to analyse the physical aspects of a device and also the extent to which this relates sensibly to the digital aspects.

In some ways this is similar to architectural models, from Seeheim onwards [PfH85], that separate presentation from functionality, but in these models the different levels are all principally digital, with only Arch/Slinky making an explicit attempt to discuss the physical level of interaction [UIM92]. However, even Arch/Slinky puts physical interaction as an additional (lowest) layer whilst we shall see that physical devices embody aspects of at least dialogue-level interaction. This is also similar to work on linking models of interface and functionality (often inspired by Seeheim), for example Moher et al's Petri Net-based 'bridging framework' for linking models of devices, users, and interfaces [MDB96]. However, these, to our knowledge, all stop short of a physical-level behaviour description.

Producing this formal framework is valuable in drawing out insights about the nature of physical interaction. However, we also believe this diagramatic representation can be useful during the design process. While formal notations are often described as being for 'communication' or 'understanding', their value in this respect is rarely subject to empirical studies (with exceptions [Joh96]). Certainly the temptation for a formalist is to add more and more features to their notation in order to make it representationally complete, but often at the expense of clarity. What might have started as a simple notation, often becomes obtuse to all but the notation's developers. Therefore, in order to understand the potential practical benefit of the approach, this paper also looks at how real designers are able to understand and produce the diagramatic notation. This is not a rigorous evaluation of the notation in practice; however, as a case study of use, it yields interesting insights into both potential features that may be needed and aspects of the notation that should not change.

This paper focuses on the physical behaviour of devices. However, there are other important aspects of the physical nature of a device including the detailed physical form and layout of the device, its ergonomic aspects, its aesthetics, and its disposition in the environment. However, we have found sufficient insight from initially restricting our scope to device behaviour, especially as this is also the level of representation that is most similar to existing user interface specification. We will revisit this issue and look forward to more comprehensive modelling at the end of this paper.

In summary our *focus* in this paper is the interactive behaviour of physical devices ‘unplugged’. Our *scope* includes models both of this physical behaviour and of the digital behaviour controlled by the device, and also the link between the two. The *goals* of the work are:

understanding: to gain generic insights into the nature of physical interaction. Following the pattern common in formal specification of user interfaces, these insights will often have informal expression once they are identified.

design: to allow designers to specify the physical aspects of novel devices and so help them design more effective interaction.

In the remainder of this section we unpack the way in which physical devices relate to their digital effects on a system’s logical state and the kinds of feedback that occur. Section 2 then reviews some critical related work. Section 3 introduces the modelling approach that is then used in Sects. 4–7, which work step-by-step through a number of example devices and systems of increasing complexity. Through these examples, the paper builds up ways of describing the devices in terms of state diagrams of both the underlying logical functions and also the physical aspects of the device. The state diagrams of the physical aspects (the device ‘unplugged’) are augmented to capture some of the interesting aspects of physical interaction (the *physigrams*). A formal model is incrementally developed that gives more precise semantics to both kinds of diagram and the relationship between them. Having developed the diagrams and formalism to deal with different kinds of existing consumer devices, Sect. 8 describes how two product designers used and adapted the physigrams as part of their exploration of alternative designs for a novel device. Finally, we reflect on the lessons learnt and further work required to obtain a complete model of physical interaction with digital devices that is both formally sound and practically useful.

1.3. Physical devices and feedback

When we use the term physical device in this paper, we are using the word slightly differently than is common. We use it to mean the actual physical button, knob or other controls on their own. For example, a light switch has properties when torn from the wall and unwired—that is the physical device is a switch whether or not it is connected to a light. In the case of a mobile phone think of the phone with its innards removed—you can hold it, press its buttons, etc., whether or not you get any digital feedback. Sometimes the term ‘physical device’ would be used for the phone together with its digital functionality, but by separating the two we aim to understand better the relationship between them.

Feedback is a critical aspect of interaction, both with digital entities and with the physical world, and plays a major role in the theory and practice of usability: effective feedback was one of Shneiderman’s principles of direct manipulation [Shn83] and one of Nielsen’s heuristics [NiM94]; also a substantial issue in the early formal modelling of interactive systems was the specification of various forms of observability [Dix91].

Once we think of the physical device and the digital effects separately, we can look at different ways in which users get feedback from their actions. Consider a mouse button: you feel the button go down, but also see an icon highlight on screen.

Figure 1 shows some of these feedback loops. Unless the user is implanted with a brain-reading device, all interactions with the machine start with some physical action (a). This could include making sounds, but here we will focus on bodily actions such as turning a knob, pressing a button, dragging a mouse. In many cases this physical action will have an effect on the device: the mouse button goes down, or the knob rotates and this gives rise to the most direct physical feedback loop (A) where you feel the movement (c) or see the effect on the physical device (b).

In order for there to be any digital effect on the underlying logical system the changes effected on the device through the user’s physical actions must be sensed (i). For example, a key press causes an electrical connection detected by the keyboard controller. This may give rise to a very immediate feedback associated with the device; for example, a simulated key click or an indicator light on an on/off switch (ii). In some cases this immediate loop (B) may be indistinguishable from actual physical feedback from the device (e.g. force feedback as in the BMW iDrive); in other cases, such as the on/off indicator light, it is clearly not a physical effect, but still proximity in space and immediacy of effect may make it feel like part of the device.

Where the user is not aware of the difference between the feedback intrinsic to the physical device and simulated feedback, we may regard this aspect of loop (B) as part of ‘the device’ and indistinguishable from (A). However, one has to be careful that this really is both instantaneous and reliable. For example, one of the authors often mistypes on his multi-tap mobile phone hitting four instead of three taps for letters such as ‘c’ or ‘i’. After some experimentation it became obvious this was because there was a short delay (a fraction of a second) between pressing a key and the simulated keyclick. The delayed aural feedback was clearly more salient than the

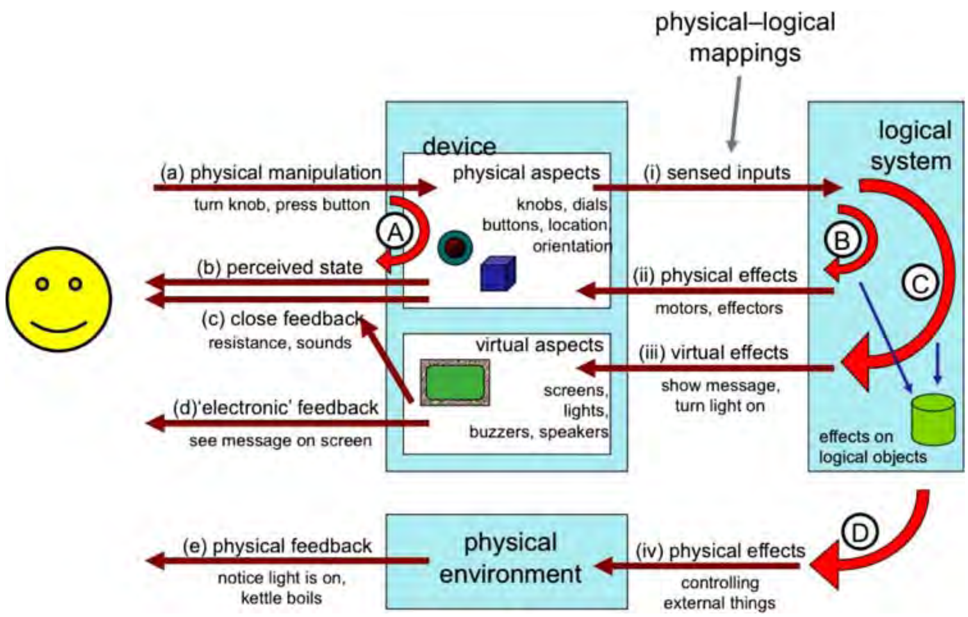


Fig. 1. Multiple feedback loops

felt physical feedback and so interfered with the typing; effectively counting clicks rather than presses. Switching the phone to silent significantly reduced typing errors!

The sensed input (i) will also cause internal effects on the logical system, changing internal state of logical objects; for a GUI interface this may be changed text, for an MP3 player a new track or increased volume. This change to the logical state then often causes a virtual effect (iii) on a visual or audible display; for example an LCD showing the track number (iii). When the user perceives these changes (d) we get a semantic feedback loop (C). In direct manipulation systems the aim is to make this loop so rapid that it feels just like a physical action on the virtual objects.

Finally, some systems affect the physical environment in more radical ways than changing screen content. For example, a washing machine starts to fill with water, or a light goes on. In addition there may be unintended physical feedback, for example, a disk starting up. These physical effects (iv) may then be perceived by the user (e) giving additional semantic feedback and so setting up a fourth feedback loop (D).

For the purposes of this paper we will not care much whether the final semantic effect and feedback is virtual (loop (C)) or physical (loop (D)), nor whether it is deliberate or accidental, as it is the physical device (that is loops (A) and if indistinguishable (B)) in which we are most interested.

2. Related work

2.1. Theories and frameworks for physical interaction

The most obvious connection to this work is Gibson's concept of affordances [Gib86]. For a simple physical object, such as a cup, there is no separate logical state and simple affordances are about the physical manipulations that are possible ((a) in Fig. 1) and the level to which these are understood by the user: Norman's 'real' and perceived affordances [Nor99]. For a more complex, mediated interface the effect on the logical state becomes critical: the speaker dial affords turning but at another level affords changing the volume. Hartson [Har03] introduces a rich vocabulary of different kinds of affordances to deal with some of these mediated interactions. In Sect. 4.2, we will look in detail at Gavers notion of sequential affordance [Gav91].

The 'Sensible, Sensable and Desirable' (SSD) framework [BSK03] deals with this relationship between the physical device and logical state. It considers three aspects of the relationship: sensable—the aspects of the physical device can be sensed or monitored by the system; sensible—the actions that the user might reasonably do to the device; and desirable—the attributes and functionality of the logical system that the user might need to control. This is used to explore the design space and mismatches between the sensible, sensable and desirable may suggest options for re-design. In Fig. 1, the sensable aspects correspond to (i), whilst the sensible ones refer to possible actions ('real' affordances) of the device (a) that the user might reasonably perform. The desirable part of the framework refers to the internal possibilities of the logical state. Note that what is sensible to do with a device depends partly on perceived affordances and partly on the user's mental model of the device and its associated logical state.

The concept of fluidity, introduced in Dix et al. [DFA04] and expanded in our work leading to this paper, is focused on the way in which this mapping is naturally related to the physical properties of the device. Whereas the SSD framework is primarily concerned with what it is *possible* to achieve, fluidity is focused on what is *natural* to achieve. This naturalness involves both cognitive aspects and also more motor-level responses. In this paper we do not explicitly consider the mental models formed by users, but this is implicit in some of the discussion of mappings. However, ongoing empirical work by the authors is aimed at teasing apart cognitive and bodily responses in physical interaction.

The work that is closest in concept to our own is Interaction Frogger [WDO04]. This discusses three kinds of feedback: inherent feedback, augmented feedback and functional feedback, which correspond almost exactly to the loops (A), (B) and (C) respectively. Physical feedback in the environment (loop (D)) is not explicitly described in their work, but would presumably fall under functional feedback. As well as feedback, the Frogger work looks at *feedforward* against each loop, where feedforward is, rather like Norman's perceived affordance, about the different ways a device/system can expose its action potential. Critically too, this work, like our own, is interested in what makes interaction natural and brings out particular qualities that impact this: time, location, direction, dynamics, modality and expression.

2.2. Device abstraction

The central focus of this paper is the detailed modelling of the physical behaviour of interaction devices themselves, with a secondary focus on the way this then links to digital functionality. To some extent this runs completely counter to the long-standing paradigm of device abstraction within user interface specification and construction.

The roots of device abstraction go back a long way, certainly to early graphics standards such as GKS and PHIGS, and were instrumental in allowing the development of applications independent of particular vendors and physical devices, indeed it is hard to envisage the modern GUI without the key abstraction of text input vs. pointer. This basic separation is evident today in the APIs for most user interface toolkits, for example, Java AWT has 'mouse' events even though the actual device is often a trackpad or stylus.

Card et al.'s Design Space for Input Devices [CMR90, CMR91] was influential in developing these notions beyond simple 2D pointers. Their analysis is particularly relevant to our work as they used as running example a radio with knobs and dials—defining these and general input devices in terms of primitive dimensions, many of which have become common language: relative/absolute, position/force, linear/rotary. Their work is also interesting in that, on the one hand, it abstracted devices into classes, but, on the other, it took into account that setting a value through rotating a dial (rotary) is different from moving a slider (linear)—that is, at least certain aspects of the physical nature of the device are important.

Device abstraction has clearly been vital to the ongoing development of interface software and hardware allowing software to be written once irrespective of the intended device and allowing new hardware, such as the trackpad, to be deployed without needing to modify or port existing applications. This is problematic in multi-modal interfaces, where the difference, for example, between speech or keyboard entry of text is intrinsic to the domain. But here too, there have been long standing efforts to add layers of abstraction, for example, in the PAC-AMODEUS architecture [NCo95].

While clearly valuable, the drive to device abstraction does elide differences that may be important, for example, the fact that a mouse button mounted on the top of a mouse may be difficult to hold down when lifting the mouse during long drag actions, whereas the equivalent interaction would be fine using a velocity joystick or inward facing mouse buttons (now rare!).

Recognition of the importance of these device differences can also be tracked to early days of HCI as a counter trend to device abstraction. It has long been recognised that devices that are functionally similar but physically different also differ in performance typically measured using Fitts' Law constants (e.g. reviews as far back as [Mil88] and [CMR90, CMR91]). However, these simple measures do not capture the full richness of device differences, for example, one of the authors recalls Milner's presentation of his review of input device performance [Mil88], where he showed images of arcade gamers who made significant use of the way a trackball continues to spin after being struck with the hand (sadly not reported in the paper itself!). Buxton also long argued for a richer view of devices with analysis of different kinds of children's drawing toys (such as Etch-a-Sketch) and using a simple finite state model to distinguish key difference between mouse, stylus and touch-based interaction [Bux86, Bux90]. Buxton's early work emphasised the way the lexical-level design of the physical interface can simplify syntax in interaction; a similar point is made by Usher et al. who refer to the 'digital syntax' embodied in the physical design of *token+constraint* tangible user interfaces [UIJ05].

More recently the importance of the physical nature of devices has re-emerged in several areas. In research in tangible user interfaces, the precise form of tangible tokens is usually designed taking into account the specific application and sensing technology [Ish08]. Also as user-interaction design has begun to overlap with product design the importance of physical form has become essential [BoV90] .

2.3. Modelling physical and continuous action

Formal modelling in human–computer interaction dates back over 25 years (e.g. [Rei81, Suf82, DiR85, ThH90, PaP97]). Most is focused at the level of the dialogue starting at symbolic actions such as 'key A pressed' and on the behaviour of the digital system; although some work includes models of the physical systems being controlled by the digital device and even the user's mental states or behaviour so that conjoint properties can be investigated (e.g. [YGS89, CuR07]). Modelling of the physical aspects of interaction devices seems rare, a notable exception is Thimbleby's recent work modelling of the layout of controls [Thi07]. This work builds on long standing finite-state models of consumer devices, such as a VCR, and augments the FSM model of the digital behaviour with a specification of the precise physical location of buttons allowing detailed timings to be estimated for multiple button presses using Fitts' Law.

To date, there are only a few formal approaches to ubiquitous and tangible interaction. The ASUR framework [DSG02, DuG07a] focuses on the arrangement of devices, people and software components in the environment and the flows of data between them. ASUR has been applied to systems embedded into the environment as well as more self-contained devices including a fairly complex bespoke device, GE-Stick, for interacting with Google earth [DGR07b]. Building on roots in multi-modal systems, the Mixed Interaction Model also deals with the structure and flows between sensors and actuators within hybrid physical–digital devices [CoN06, CoN08], and has been applied to the design of a hand-held photo browser, not unlike that in Sect. 8 of this paper. Like ASUR, the approach is focused principally at the flows and relationships between sensors, but also includes tool support down to concrete implementation. At a more detailed level the uppaal system, which is based on timed automata, has been used to model a navigation system to guide visitors in an office building [HKC07]. While addressing a ubiquitous application, the features modelled in uppaal are principally those of the information system with the knowledge of the physical layout of the building embodied in a single black-box function '`sensorlink`()' giving the closest navigation display to a user's location.

One of the crucial differences between the physical and digital worlds is *continuity* in terms of space, value and time. In this paper, we wish to represent simple phenomena as simply as possible, and so we will largely avoid modelling continuous activity and values. However, we shall find that the intrinsic continuity of physical world asserts itself, and issues such as in-between positions of switches, muscle pressure, and timed behaviour have to be considered.

To our knowledge, the earliest work that systematically addresses issues of continuity is status–event analysis. This includes formal models, specification notations and a conceptual vocabulary for dealing with such systems [Dix91b, DiA96]. Most recently this has been developed into an XML-based implementation notation XSED [DLF07]. Status–event analysis distinguishes event phenomena such as the (moment of the) pressing of a button, from status phenomena, such as the position of a mouse or the current temperature. Going further back, this distinction is implicit in the difference between event from sampling devices in GKS and very early formal device models [Ans92].

Most specification and modelling techniques for continuous interaction use two linked specifications largely separating the continuous and discrete parts. The discrete part is some form of discrete state system with event transitions whereas the continuous part defines status–status mappings that relate continuous input and output phenomena. The two are linked in that critical changes in status phenomena are treated as events (status-change events) and the exact form of the status–status mapping depends on the current state of the discrete system.

Two early examples of this are Interaction Object Graphs [Car94], which used a variant of Harel State charts for the discrete part augmented with data flows to capture the continuous interactions; and the PMIW user interface management system, which managed status–status mappings using a data-flow notation, which is effectively 'rewired' by a discrete finite state machine driven by event inputs [JDM99]. The former was used to model novel on-screen widgets and the latter to model interactions in virtual environments such as grabbing and moving objects.

Various forms of Petri Net have also been used to specify virtual environments [MDS99, WiH00]. In particular, Smith has recently used Flownets to model haptic interaction in virtual environments [Smi06]. As in Interaction Object Graphs and PMIW, Flownets use data flows to model the continuous (status–status) aspects of the system, with Petri Nets for the discrete aspects. The two are linked with the data flows being able to initiate tokens into the Petri Nets through threshold triggers (status-change events) and the presence of tokens at particular points being able to regulate the dataflow using 'flow controls'.

An exception to the use of largely separate discrete and continuous components was Wüthrich's work using cybernetic theory to model both discrete and continuous parts within what is effectively a purely continuous paradigm [Wüt99]. Within this paradigm discrete events become continuous functions that are zero/undefined except at the exact moment when the event occurs. This use of continuous techniques originating from a physics/engineering background is rare, however recent work by Eslambolchilar has used control theory to model human control of interface elements as a single (typically closed-loop feedback) system, which has been applied to a number of screen-based controls and mobile devices [E06].

3. Modelling approach

In Sects. 4–7 we will examine a number of properties of physical device interaction that have been identified from our previous studies of consumer products [GhD03, GhD05]:

- exposed states
- bounce back
- time dependent devices
- controlled state and compliant interaction.

This is not the complete list of properties uncovered in previous analysis, but includes those where the more formal analysis of this paper adds insight, or where the nature of the property requires additional elements in the physigrams. For example, another property is the 'natural inverse': whether physical movements that are naturally opposite (push/pull, etc.) cause opposite system changes. While the natural inverse is important and is currently being studied in detail, it is most centrally about the physical layout of controls and how these link to human motor movements, and so lies outside the scope of this paper. The natural inverse has interesting parallels with Task-Action Grammar [PaG86], which was specifically formulated in order to deal with natural relationships in textual interaction (e.g. a command 'U' for 'up' should be matched with 'D' for 'down'). This suggests that adding a model of physical form could add significant value to physigrams; however, for this paper we focus on behaviour only.

Each of the sections will start with an informal example of consumer devices exhibiting the property, from a light switch to a washing machine. The example is then specified using an (augmented) STN of the physical device itself (the physigram) and an associated STN of the underlying logical state. As the sections unfold extra features are added to the physigram. Alongside the development of the physigrams themselves, each section also includes development of a more abstract formal model that directly encodes the elements in the physigram and logical state STN, and allows the specification of properties of them and their relationship. The formal model is itself grounded by showing the example systems described in the developing model.

The formal model gives what could be regarded as a surface semantics to the diagrammatic examples; that is it directly models the elements in the diagrams, but does not relate them to any independent semantic framework such as a physical model of the world. This is of course the normal level of semantic description in interface specification and indeed formal specification in general. However, the fact that we are focused on physical devices does suggest that some deeper semantics would also be of value, not for day-to-day use, but in order to give stronger foundations and in order to link different kinds of models.

The physigrams are basically slightly augmented STNs applied to the physical device. There were several reasons for using STNs rather than some other formalism.

First is simplicity. Only one of the authors is a formalist and in particular the author who performed the majority of the analysis of existing devices does not have any training or experience in formal notations or analysis. The physigrams have therefore been developed to support the needs of those without a formal background. The comprehensibilty of STNs is evidenced by the fact that they are used in end-user documentation; for example, an STN used in digital watch documentation is reproduced in the 'dialog notations and design' chapter of [DFA04]. Because STNs are relatively easy for non-experts to interpret, they are also used to communicate user interface issues to broad or popular audiences, for example, Degani uses STNs extensively in 'Taming Hal' [Deg04]. In fact, Degani includes an illustrative example where an STN is used to specify a light switch (p. 14) in a way that looks very close to the early examples of physigrams here, but does not distinguish the switch from the light it controls.

This does not mean that STNs are without problems, and we are aware that it is far more difficult to create STNs than to read them; in particular our experience using user interface STNs with students and at tutorials has been that novices find it difficult to distinguish activities/events from states. Interestingly this did not seem to be a problem with any of the physigrams produced by designers in Sect. 8; possibly this is because for a physical device the states are far more apparent than for an user interface where states may be hidden or the appropriate level of abstraction unclear.

As well as being relatively simple, STNs are actually quite powerful and variants of state transitions have been used for user interface specification from Parnas in 1969 [Pa69] to Thimbleby's long term body of work looking at automated analysis of consumer user interfaces and his recent book 'Press On' [Thi07]. Indeed, many of the alternative formalisms are either variants of STNs or can be rewritten as STNs for practical examples. The most obvious alternative would be statecharts [Har87] as used in UML and used also for more complex examples in both Degani and Thimbleby's books [Deg04, Thi07]. The parallel behaviour of the device STN and logical system STN could be described in a single statechart, but this did not seem to add additional expressivity beyond juxtaposing the STNs informally.

Finally, we use STNs because their simplicity means they embody fewer assumptions/biases than more complex notations, which, by their nature, tend to overcommit—especially when attempting to specify continuous behaviour in finite notations [DiA96b]. This will become evident in Sect. 7 when we discuss the way system events may change device states in a way that may 'fight' with user actions. If, for example, we had used statecharts to model the link between physical device and logical state, this would have included a default semantics for synchronisation, which would not have actually expressed the physical situation.

In general we have tried to avoid shoehorning the example devices into a pre-defined notation and instead attempt to flexibly change the notation to express the problems and features clearly and with verisimilitude. Having done this, it would be possible to look at each of the new features and ask how they would be modelled in, or be added to other notations, such as statecharts or timed Petri nets, although always doing so with an awareness of the intended audience of the resulting notation.

4. Exposed states and physical–logical mapping

4.1. Example: up/down light switch

One of the simplest examples of a physical device is a simple on/off light switch. In this case the switch has exactly two states (up and down) and pressing the switch changes the state (Fig. 2a).

Actually even this is not that simple, as the kind of press you give the switch depends on whether it is up and you want to press it down or down and you want to press it up. For most switches you will not even be aware of this difference because it is obvious which way to press the switch . . . it is obvious because the current state of the switch is immediately visible.

Note that the switch has a perceivable up/down state whether or not it is actually connected to a light and whether or not the light works.

The logical system being controlled by the device also has states and Fig. 2b shows these in the case of the light bulb: simply on or off. (In fact the light bulb may also be broken, but we will ignore faults in this paper.)

Of course in the case of a simple light switch, the states of the physical device are in a one-to-one mapping with those of the logical system being controlled. In previous work we have used the term *exposed state* [GhD03] to refer to the way that the perceivable state of the device becomes a surrogate for the logical state and makes it also immediately perceivable. In the case of turning on an incandescent light bulb in the same room as the light switch, this is a moot point as the semantic feedback itself is immediate and direct. However, in some cases there may be a delay in the semantic response (e.g. neon lights starting up, kettle when first turned on) or it may be hidden (e.g. external security lighting); in these cases the feedback inherent in the device is not just very obvious, but may be the only immediate feedback.

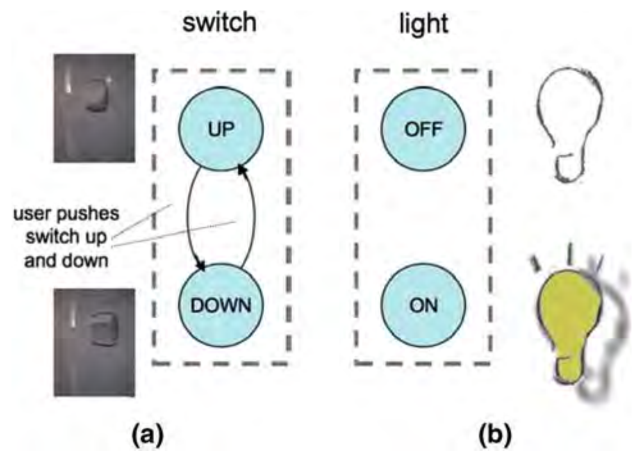


Fig. 2. Light switch: **a** physical device, **b** logical states

4.2. Formal model

We can model this kind of behaviour more generically. We denote by UA the set of potential user actions such as ‘push up’; these may be particular to a specific device ‘push button A’ as our environment affects our action possibilities. We use PA to denote the set of perceivable attributes of the world ‘light is shining’, ‘switch is up’. The full perceivable state of the world is composed of the different perceivable effects and there may be masking effects, e.g. if light 1 is on we may not be able to tell that light 2 is also on. However, for simplification we will just assume these are individually identifiable—at least potentially perceivable.

The physical device we model as a simple state transition network:

DS − physical states of device
$DT \subseteq DS \times DS$ − possible device transitions

In the light switch every transition (only two!) is possible, but in some situations this may not be the case. Hence the physically possible transitions are a subset of all conceivable from–to pairs. Note, for brevity, we will assume that the physical device states DS consist solely of those reachable through physically realisable transitions DT. For example, if the device consisted of two seamless hollow balls, we would not include the state where the smaller ball ‘magically’ was inside the larger.

Some of these transitions are controlled by user actions:

$action\colon UA \leftrightarrow DT$ −n − m partial relation

Note that this relation is n-to-m, that is the same user action may have an effect in several states (with different effect) and a single transition may be caused by several possible user actions (e.g. pressing light switch with left or right hand). In addition neither side is surjective, some physically possible transitions may not be directly controllable by the user (e.g. lifting large weight, pulling out a push-in switch) and some user actions may have no effect in the device in any state (e.g. blowing your nose). However, for exposed-state devices we will normally expect that the states are completely controllable by the user within the physical constraints of the device:

$controllable\text{-}state \equiv action$ is surjective

Aspects of the user’s state may be perceivable by the user:

$ddisp\colon DS \rightarrow PA$

And in the case of exposed state each device state is uniquely identifiable by its visible or other perceivable attributes:

$exposed\text{-}device\text{-}state \equiv ddisp$ is injective

Finally the logical system also has states which themselves may be perceivable via the feedback loops C or D.

LS — logical states of system
$ldisp : LS \to PA$

For any system we can define a map describing which device states and logical states can occur together:

$state\text{-}mapping : DS \leftrightarrow LS$

The precise nature of this mapping depends on the operation of the system. In some cases like the light switch this is a one-to-one mapping between the physical device and logical states and this is precisely what we mean by *exposed state*.

$exposed\text{-}state \equiv state\text{-}mapping$ is one-to-one

This concept of exposed state is similar to some of the strongest forms of *observability* in the early formal methods in HCI literature [DiR85, Dix91]: the internal state of the digital system is completely observable through the external appearance or other physical properties of the device. However, exposed state is stronger than these observability properties as the means to manipulate the state are exactly those through which the state is observed. Looking back to this early literature, the system is also completely *predictable* so long as the user understands (a) the manipulations possible on the device and (b) the mapping between device states and system states.

4.2.1. *Modelling the example*

The mapping between the diagram components and the model is direct. As an example, we can express the light switch from Fig. 2 as follows:

$DS = \{\texttt{UP}, \texttt{DOWN}\}$
$DT = \{< \texttt{UP}, \texttt{DOWN} >, < \texttt{DOWN}, \texttt{UP} >\}$
$UA = \{\texttt{PUSH\_UP}, \texttt{PUSH\_DOWN}\}$
$action = \{\texttt{PUSH\_DOWN} \mapsto < \texttt{UP}, \texttt{DOWN} >, \texttt{PUSH\_UP} \mapsto < \texttt{DOWN}, \texttt{UP} >\}$
$PA = \{$ , , ,  $\}$
$ddisp = \{\texttt{UP} \mapsto$ , $\texttt{DOWN} \mapsto$  $\}$
$LS = \{\texttt{OFF}, \texttt{ON}\}$
$state\text{-}mapping = \{\texttt{UP} \mapsto \texttt{OFF}, \texttt{DOWN} \mapsto \texttt{ON}\}$
$ldisp = \{\texttt{OFF} \mapsto$, $\texttt{ON} \mapsto$  $\}$

Note that *ddisp* is injective so it is an *exposed-device-state* device and also *state-mapping* is one-to-one so the system has *exposed-state*.

4.3. Physical constraints as dialogue

The physical nature of a device puts limits on what can be achieved with it. We made DT a subset of conceivable transitions because some potential transitions between states may not be physically realisable; for example, a dial may not be able to move from setting 1–3 without first going through setting 2. Also we have discussed how there may be states that are conceivable, but cannot be achieved through reasonable physical transitions . . . at least not without physically breaking or dismantling the device.

In traditional UIMS and user interface architecture literature, one of the central concepts is dialogue—the component(s) responsible for the order and interpretation of interaction depending on context. However, physical interaction is usually placed at the lowest levels, separated from the dialogue by an intermediate layer (presentation in Seeheim [PfH85], lexical/logical interaction in Arch–Slinky [UIM92], and both presentation and interaction toolkit components in PAC-Amodeus [NCo91]). This is largely because the assumption underlying these architectures is that the physical devices are generic and do not have any constraints on their interaction—on a keyboard any key may be pressed at any time. With such an assumption, constraints on the user's possible interactions are governed by the way in which the unrestricted physical interactions of the user are interpreted by the system—dialogue imposed by software. In contrast more specialised devices may impose constraints on

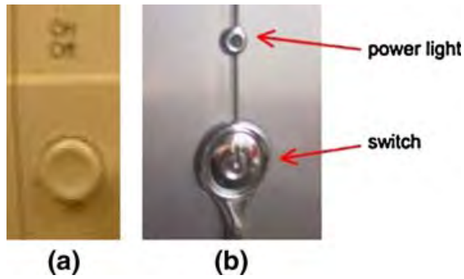


Fig. 3. a On/Off control with bounce back—is it on or off now? **b** On/Off button with indicator light

the possible interactions by virtue of their physical design; precisely the issue identified in Buxton's early work on simplfying dialogue syntax through physical design [Bux86, Bux90] and Usher et al.'s concept of 'digital syntax' [UIJ05]. As an example consider a mobile phone where the keyboard needs to slide open (which as a side effect causes a software event) before it can be seen and touched . . . and hence used—dialogue imposed by hardware.

In fact, discussion of this form of physical constraint can be found in some of the early work on formal modelling of user interfaces in the concept of a 'language' for the PIE model [Dix91]. The example used in this early work was a cash dispenser or ATM, as many ATMs at that time had a shield that covered the keypad and screen; only when a bank card was inserted was it possible to type at the keypad. These physical constraints can be seen as imposing a form of dialogue. For example, the old ATM effectively had a (low-level) dialogue of the form:

```
ATM ::= Transaction*
Transaction ::= CardIn  [0-9]*  CardOut
```

Note too that CardIn and CardOut must alternate—you cannot remove the bank card unless it is in the machine! This constraint is even true of generic keys and mouse buttons where 'press' and 'release' events must alternate. Physical interlocks are also common in machinery and industrial plant where the consequences of doing things out of proper order can be catastrophic.

This restriction of dialogue through constraints on possible interactions is not only found in 'real' physical interfaces. Even in WIMP interfaces a dialogue box, menu or window must be displayed on screen in order for the user to interact with it. It is often not necessary to have a software rule that says 'user is not allowed to press button A unless condition C holds' as one simply makes sure that button A is not visible on screen. This embedding of dialogue into the virtual 'physical' constraint of what is available on-screen, is probably the reason for the low emphasis on *explicit* dialogue management in much GUI interface construction.

For the designer of digital devices, the embedding of what would otherwise have to be software dialogue into physical constraints, can be an opportunity to both reduce the complexity of the digital interaction and make the interaction more intuitive.

5. Bounce back buttons

5.1. Example: push on/off switch

A more complex behaviour occurs with bounce-back buttons or other devices where there is some form of unstable state (pressed in, twisted) and where you need to exert continuous pressure in order to maintain the state. Figure 3a shows a typical example of a computer on/off switch. One press and release turns it on, a second turns it off.

Note that the user action here, pressing the button, is not discrete, but involves pressing until it 'gives' then *releasing*. While you maintain pressure the button stays in, it is when you release that it comes out again. However, the button comes out not because you pull it out with your release of pressure, but because it is internally sprung—the bounce-back.

Bounce-back buttons are found everywhere, from keyboards, to mice, to television sets. They typically do not expose the state of the underlying system as there is just one stable state of the physical device and it is the history of dynamic interactions with it that is important (on or off, what channel). The temporary unstable states of the device are distinguishable, so long as pressure is maintained, but the device returns to its stable state as soon as the pressure is released. That is, the physical device itself does not maintain a record of its interaction in the way a rocker switch does.

Because the device does not itself expose the state of the underlying system (there is no feedback loop A for the state) we get potential problems of *hidden state* [GhD03]. Sometimes this is not an issue because the semantic feedback (loop C or D) is sufficient—for example, switching channels on a television set. However, even where there is semantic feedback this may be ambiguous (switching channels during an advertisement break) or delayed (the period while a computer starts to boot, but is not yet showing things on screen). In such cases, a supplemental feedback (loop B) close to the device is often used, such as a power light on or near the switch (as in Fig 3b).

Where the device does not have any intrinsic perceptible tactile or audible feedback (e.g. click of the switch or feeling of 'give' as it goes in) then supplemental loop (B) feedback may be given for the transitions as well as the states. Simulated key clicks or other sounds are common, but also, more occasionally, simulated tactile feedback can be used, as in the BMW iDrive.

From a design consideration, indirect feedback, whilst less effective, is useful in several situations:

- Where the complexity of the underlying system exceeds the potential states of the device, a simple one-to-one mapping is not possible. For example, a channel dial could work in a one-to-one mode if there are half-a-dozen channels, but not if there are hundreds of channels to choose from.
- Where we want to generate something that is perceived of as an event or action on the system. For example, the big red button (beloved of B movies) that fires a missile; here a rocker switch would make no sense, you can't 'unfire' the missile by pushing the switch back.
- Some logical state transitions may be under internal system control. For example, a computer may be turned on using a push button, but switching off may be 'soft', under the control of the computer, to prevent accidental data loss. In Sect. 7 we will return to this issue and see how exposed-state can be consistent with system control. It should be noted that neither of the computer on/off switches photographed in Fig. 3 are on computers which can power themselves down in this way.
- The fact that continuous pressure is required can be used explicitly in tension states [Dix91] in order to manage temporary modes. Modes in interfaces have long been known to be a problem; the meaning of an action depends on the mode and so the effect of an action may not be as intended if the user does not know the current mode. Additional feedback is often added to make modes perceptually obvious most frequently visually (e.g. cursor shape), but also aurally [Mon86]. Associating modes with tension states mean that there is *haptic* feedback so it is hard to 'forget' that you are in a mode. For example, moving the mouse with a button pressed might draw a line rather than simply move the cursor, but users do not confuse the two as they can feel their finger pressing the button.
- A special case of tension state modes is when there is some sort of time-dependent effect in a mode (e.g. velocity-base joysticks). We will discuss time-dependent devices in Sect. 6.

5.2. Formal model

We can inherit much of the same formal machinery developed for simple exposed-state devices. However, in addition to transitions controlled by the user, we have bounce-back transitions. We label them Z (after Zebedee and the fact that Z and S are used for left- and right-handed helices). In the example here there is only one action leading to the states and thus it is clear what kind of user tension needs to be released in order for the bounce-back to happen. Sometimes (and we will see an example later) there is more than one tension simultaneously for a state (e.g. twist and pull), so we need to label the bounce-backs by the kind of release of tension (in terms of user action) that is being released.

$$Z : UA \leftrightarrow DT$$

The states that are the subject of bounce-back transitions are transitory states for that user action:

$$\forall a \in UA \ \textit{transitory-states}(a) \equiv \{d \in DS \ st. \ \exists (d, d') \in Z(a)\}$$

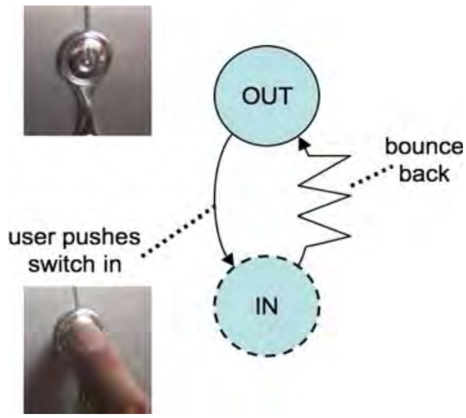


Fig. 4. States of bounce-back button

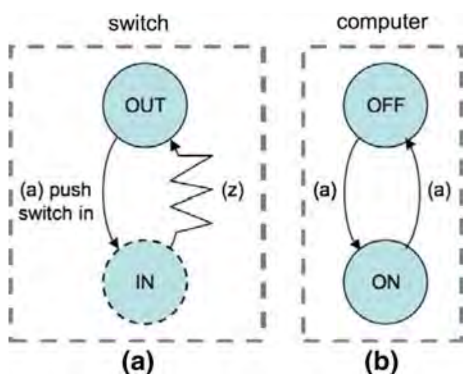


Fig. 5. **a** Physical states changes trigger event (a), **b** logical state changes based on events

Furthermore a transitory state for a user action cannot be the source of the same user-controlled transition and must have been reached by that user action:

$$\forall a \in UA\ \textit{transitory-states}(a) \cap dom(action(a)) = \{\} \\ \wedge\ \textit{transitory-states}(a) \subseteq range(action(a))$$

Figure 4 shows the example of the computer switch with the bounce-back transition shown as a zig-zag line (spring) and the transitory state (IN) dotted.

While exposed state devices can have a one-to-one mapping between logical states and physical states, here the relationship is based on the events. Formally we define this first by associating events from a set Ev with physical state transitions:

$$trigger : DT \rightarrow Ev$$

This mapping may be partial as not every transition will cause an event. Also it is typically the case that only user-controlled transitions cause events ($dom(trigger) \subseteq range(action)$), because once you have pressed a switch you are committed. However, there are exceptions such as the ‘drop’ (release the button) when you drag and drop with a mouse.

These events then cause transitions in the logical system:

$$doit : Ev \times LS \rightarrow LS$$

Figure 5 shows the physical device STN annotated with an event (a) and the effect of the event on the logical state (computer power). Note that in this example (and it is common!) there is no reason why the system could not have been designed with exposed state, for example a button that stays depressed and requires an extra push to release it. This design choice is often motivated by the aim to have a smooth surface although in the example in Fig. 3b the switch is part of an embellishment anyway, so even this aesthetic reason seems to be absent. Finally, one of the reasons listed for having bounce-back is when the system is able to power itself down (or some similar logical state transformation), however, as noted previously, this is not the case for the switch in Fig. 3b.

5.2.1. Modelling the example

As in the previous section we can apply this to Figs. 4 and 5:

$DS = \{\texttt{OUT}, \texttt{IN}\}$
$DT = \{< \texttt{OUT}, \texttt{IN} >, < \texttt{IN}, \texttt{OUT} >\}$
$UA = \{\texttt{PRESS}\}$
$action = \{\texttt{PRESS} \mapsto < \texttt{OUT}, \texttt{IN} >\}$
$Z = \{\texttt{PRESS} \mapsto < \texttt{IN}, \texttt{OUT} >\}$
$transitory\text{-}states(\texttt{PRESS}) = \{\texttt{IN}\}$
$PA = \{$ , , $\ldots\}$
$ddisp = \{\texttt{OUT} \mapsto$ , $\texttt{IN} \mapsto$  $\}$
$LS = \{\texttt{OFF}, \texttt{ON}\}$
$Ev = \{\texttt{(a)}\}$
$trigger = \{< \texttt{OUT}, \texttt{IN} > \mapsto \texttt{(a)}\}$
$doit = \{< \texttt{(a)}, \texttt{OFF} > \mapsto \texttt{ON}, < \texttt{(a)}, \texttt{ON} > \mapsto \texttt{OFF}\}$
$state\text{-}mapping = \{\texttt{OUT} \mapsto \texttt{OFF}, \texttt{OUT} \mapsto \texttt{ON}, \texttt{IN} \mapsto \texttt{ON}\}$

We can verify the conditions on *transitory-states*:

$$transitory\text{-}states(\texttt{PRESS}) \cap dom(action((\texttt{PRESS}))) = \{\texttt{IN}\} \cap \{\texttt{OUT}\} = \{\}$$
$$transitory\text{-}states(\texttt{PRESS}) = \{\texttt{IN}\} = range(action((\texttt{PRESS})))$$

Looking at $ddisp$, it is injective, so, like the switch in Fig. 3, this too is an *exposed-device-state* device; the IN and (transitory) OUT states are distinguishable. However *state-mapping* is not one-to-one so the system does not have *exposed-state*.

5.3. Recapitulation: the exposed state switch

Using this expression of push-back we could in principle use this to model in greater detail the exposed state switch capturing the fact that pressure has to be initially exerted and some slight give is felt until the switch eventually yields and flips to the new state. Figure 6a shows this with transitory states for when the switch is up and just being pushed down. If you release before putting sufficient pressure on it snaps back to UP, but if the pressure is sufficient the switch yields and goes to the new state.

This yielding is rather like bounce back in that once the critical point is reached the device just goes of its own accord. However, we have drawn it slightly differently (less of a spring and more of a lightning bolt) in order to emphasise that this is going 'with' the user's action and it is the point at which the 'commitment' occurs.

Note that in Fig. 6a a transition is included for 'press up' in the UP state which simply leaves the switch in the UP state. This distinguishes 'press down', for which there is a little give with a small pressure, from 'press up', for which there is no give. Thus we can begin to capture some of the nature of Gaver's sequential affordances (described below).

In fact to model this completely we would need to include degrees of pressure and the fact that there is not just one half pressed-down state, but a whole series requiring increasing pressure. This is not captured by a finite state diagram or description and would require a full status–event description as we are talking here about *interstitial behaviour* (the interaction between events) and status–status mapping (more pressed = more down) [DiA96]. This is also reminiscent of Buxton's three-state model for pointing devices, where the degree of finger pressure is one of the critical distinctions between abstract states [Bux90].

Figure 6a is rather complicated and, whilst useful in order to clarify detailed behaviour, would be a little noisy for real design use. Figure 6b shows a shorthand that emphasises the slight give of the press down action in the UP state by the comic-book-style movement arcs. In fact, we will often omit even this, as in many cases every action has this slight give property. However, in Sect. 8 we will see examples where explicitly marking give vs.

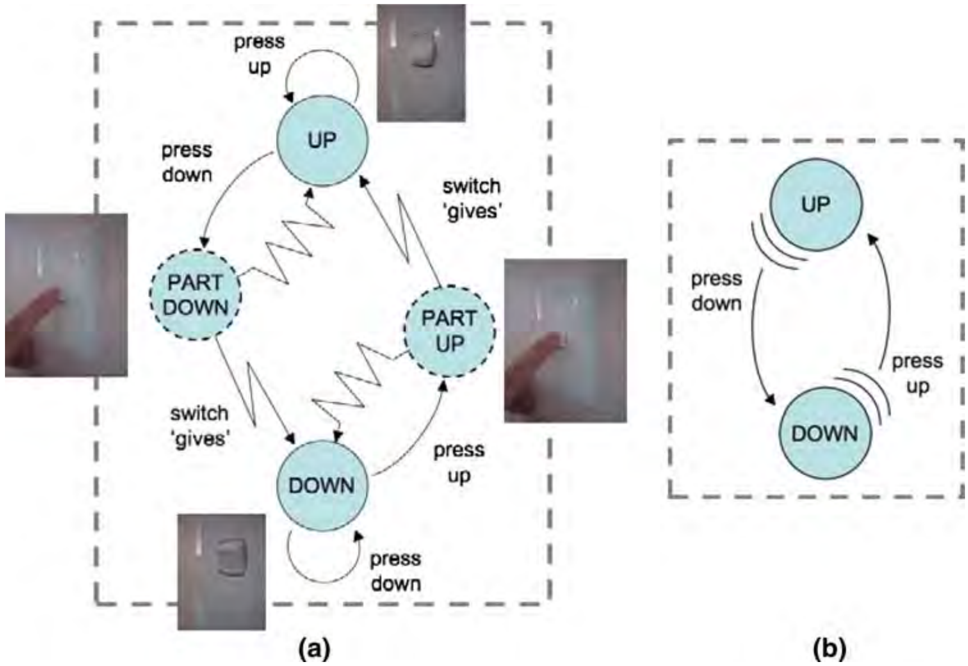


Fig. 6. Capturing initial pressure on exposed state switch: **a** detailed model using bounce-back, **b** more convenient shorthand

non-give transitions distinguishes otherwise similar physigrams. This form of shorthand would be also be useful in cases where some controls are operated on the slightest pressure—typically fully electronic ones. Formally we can capture this 'give' in some transitions by a simple '*has-give*' predicate over user actions in particular states.

5.4. Give and sequential affordance

Note how even in this simple example of flipping a switch, the user actions are not simply 'events', but are protracted over time and vary in force. An additional way in which the user gets feedback on the state of the device and appropriate actions is by 'trying out' an action, often unconsciously, and if there is a small 'give' in the device continuing the action and increasing pressure until the user action causes a change in state of the device.

The importance of this effect is hinted at by Gaver when he introduced the notion of *sequential affordances* [Gav91]. Gaver discusses a door handle which initially just looks the size of your hand, so invites grabbing. However, once you have the door handle, it then has a second affordance, that of turning. It maybe that in that early paper Gaver simply meant (in the sense of Gibson's affordances [Gib86]) that a door handle is not physically turnable by your hand until you have grapsed it. However, it is also the case that when your hand is on the door handle you can feel a little 'give' and this tells you which direction to turn the handle, especially important when it turns in the 'wrong' direction.

This use of 'give' is clearly part of tacit design practice, for example, many cameras use a half-pressed shutter release button to mean 'do auto focus'. However, to the authors knowledge, the issue is not discussed more explicitly in the HCI literature. This is perhaps surprising given the importance of the distinction between, say, a touch-based switch and one with a more solid button, but perhaps the lack of attention to such differences is simply due to the prevailing culture of device abstraction.

Indeed once the issue of 'give' is forefronted we can see digital equivilants, such as tooltips or the ability to slide off an on-screen button with activating it. This could also impact hardware design; for example, if mouse buttons had a 'pressing but not fully-pressed' state, rather like the camera shutter release, then this could be used as part of interaction to show some form of preview of the effect of the click, just like trying a door handle.

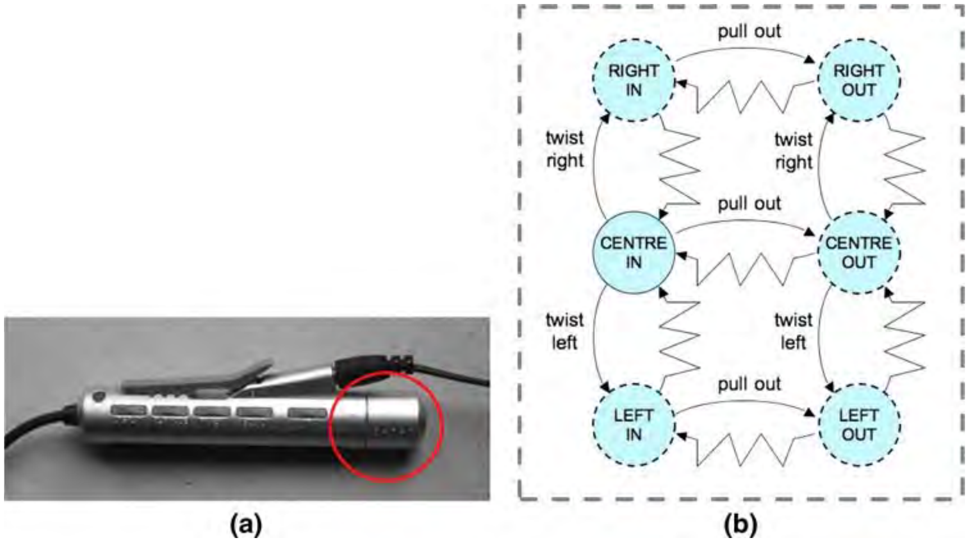


Fig. 7. a Minidisk controller; **b** device states

6. Time-dependent devices

6.1. Example: track/volume selector

Our next level of complexity includes devices, such as keyboards with auto-repeat or tuning controls on car radios, where things happen depending on how long you have been in a state. Figure 7a shows a minidisk controller. The knob at the end can be pulled in or out turned to the left or right. Figure 7b shows this physical state transition diagram of the device. This might be more succinctly described using two 'concurrent' STNs, one for in-out and one for left-right (as in state charts), but as they are coupled in a single control we are showing all the transitions to give an idea of the total complexity of even such a simple thing as one knob!

Whether the knob is pulled in or out determines whether it is affecting the volume or track selection and the amount of time it is pushed to the left or right moves the volume/track selection up or down. The former is a simple mode effect . . . and, as discussed in Sect. 5, a tension mode carries its own feedback, so is a good design feature. However, we shall focus on the left-right twists and their time behaviour.

To do this Fig. 8a shows just the left-right part of the diagram (actually in the 'out' condition) for when it is controlling track selection, and Fig. 8b shows the state diagram for the logical system, the selected track. As in Fig. 5 we use event labels to match the two. For this device we have had to augment the device transitions with additional timed transitions (labelled τ).

For some devices, there may be timed behaviour as part of the physical device itself, for example, eco-friendly light swtches in hallways that slowly turn themselves off. However, for the minidisk controller, these timed transition are not part of the physical device behaviour, but are strictly part of the device–logical state mapping; Fig. 8a is thus not the raw device STN. We have added them as *annotations* to the device STN, both for convenience, and also because the user is aware that the knob is *being held* for some time even if the exact times when events are triggered are not totally under the user's control (unless they have a millisecond clock in their heads!). In Sect. 8.3 we will again see a need to 'layer' some additional information onto the raw device physigrams, but whenever we do this we need to be very careful as we are adding information that the designer knows, but may not be apparent to the user from the physical behaviour of the device.

From a usability point of view, these timed events have a special status as the user is not performing clear actions. In the case of the minidisk controller, the timed events are all in tension states increasing the user's awareness that additional system events may occur. This follows one of the general design heuristics from status–event analysis that *trajectory dependent* effects (those where the path of movement matters, not just its end point) should normally take place only in tension states [Dix91]. A very easy 'undo' is even more critical for these implicit timed events than for more deliberate user' actions. In the case of the minidisk controller there is no explicit 'undo',

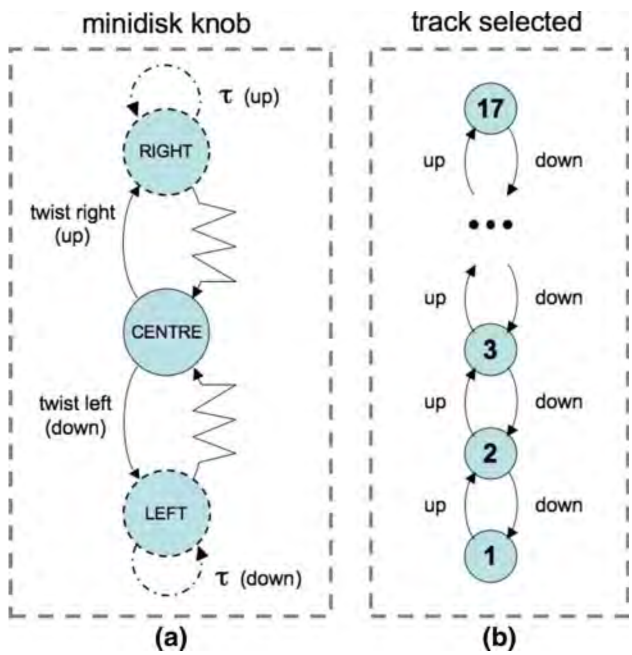


Fig. 8. Minidisk: **a** time augmented device **b** logical states

however simply turning the knob in the opposite direction creates the opposite effect. In fact, this form of *natural inverse* is a very easy way of allowing the user to effortlessly undo actions [GhD06].

In addition to timed events in or closely related to the device behaviour, there are often timing issues more deeply embedded in the digital behaviour; for example, many digital appliances revert to some default state after some period of inactivity. There is a long-standing literature on the importance of time in the user interface (e.g. see [Shn84] and chapter 5 of [Dix91]), but it is still often not given due attention in interaction design. However, given our focus on the physical device behaviour, we do not consider these more internal timings further.

6.2. Formal model

Notice that everything in Fig. 7b, with the exception of CENTRE-IN, is a tension state. However, there are actually two kinds of tension demonstrating why we needed to label transitory states and bounce-backs by user actions in Sect. 4.2.

In Fig. 8a we draw the timed events as if they were transitions, however we model them simply as an aspect of the state. This is because for the user the system does not make little transitions to the same state, it simply stays in the tension state. For the user the transitions happen *while* in a state not at some hidden transition in an invisible state model. This emphasises the importance of allowing the phenomena to shape the notation rather than fitting phenoena to the notation as discussed in Sect. 3. There may also be real timed transitions, but these are more often in response to things happening in the logical state, which we discuss in the next section. So all we need to do is say in which state and how frequently the timed events occur.

$$\textit{time-trigger} : DS \times Time \times Kind \rightarrow Ev$$

Here $Time$ is as in 'gap between moments' rather than time on the clock, and $Kind$ is either $PERIODIC$ or $SINGLE$. This is not totally general, but seems to capture most timed events seen in practice except complex continuous time effects such as mouse 'acceleration' settings. Note that in other circumstances we would be able to dispense with the special $SINGLE$ case by adding an extra state. However, the states of the STN correspond exactly to the physical states of the device, so we cannot simply duplicate them corresponding to hidden electronic or digital transitions. The only situation where it would be appropriate to add time-based states to the device

STN would be when this is apparent in the state of the actual physical device, for example, in the way that some corridor light switches turn themselves off after a fixed time.

In terms of status—event analysis, these timed events are another example of interstitial behaviour. This again shows that a more fine-grained model would need to use a full status–event description and we would need to use some form of a real-time model to express precisely the detailed semantics of *time-trigger*.

6.2.1. *Modelling the example*

The physigram for the full minidisk knob in Fig. 7 can be modelled as follows:

$$\begin{aligned}
DS &= \{\texttt{CENTRE_IN}, \texttt{LEFT_IN}, \texttt{RIGHT_IN}, \texttt{CENTRE_OUT}, \texttt{LEFT_OUT}, \texttt{RIGHT_OUT}\}\\
DT &= \{<\texttt{CENTRE_IN}, \texttt{LEFT_IN}>, <\texttt{CENTRE_IN}, \texttt{RIGHT_IN}>, <\texttt{LEFT_IN}, \texttt{LEFT_OUT}>, \ldots\}\\
UA &= \{\texttt{TWIST-LEFT}, \texttt{TWIST-RIGHT}, \texttt{PULL-OUT}\}\\
action &= \{\texttt{TWIST-LEFT} \mapsto <\texttt{CENTER_IN}, \texttt{LEFT_IN}>, \texttt{TWIST-RIGHT} \mapsto <\texttt{CENTER_IN}, \texttt{RIGHT_IN}>,\\
&\quad \texttt{TWIST-LEFT} \mapsto <\texttt{CENTER_OUT}, \texttt{LEFT_OUT}>, \texttt{TWIST-RIGHT} \mapsto <\texttt{CENTER_OUT}, \texttt{RIGHT_OUT}>,\\
&\quad \texttt{PULL-OUT} \mapsto <\texttt{LEFT_IN}, \texttt{LEFT_OUT}>, \texttt{PULL-OUT} \mapsto <\texttt{CENTER_IN}, \texttt{CENTER_OUT}>,\\
&\quad \texttt{PULL-OUT} \mapsto <\texttt{RIGHT_IN}, \texttt{RIGHT_OUT}>\}
\end{aligned}$$

$$Z = \{\texttt{TWIST-LEFT} \mapsto <\texttt{LEFT\_OUT}, \texttt{CENTER\_OUT}>, \ldots, \texttt{PULL-OUT} \mapsto <\texttt{LEFT\_OUT}, \texttt{LEFT\_IN}>, \ldots\}$$

$$\begin{aligned}
transitory\text{-}states = \{&\texttt{TWIST-LEFT} \mapsto \texttt{LEFT_IN}, \texttt{TWIST-LEFT} \mapsto \texttt{LEFT_OUT},\\
&\texttt{TWIST-RIGHT} \mapsto \texttt{RIGHT_IN} \ldots, \texttt{PULL-OUT} \mapsto \texttt{LEFT_OUT}, \ldots\}
\end{aligned}$$

Note how the *transitory-states* include several types of user action for some states. For example, LEFT_OUT requires pressure of both TWIST-LEFT and PULL-OUT. Note also how Z records which state LEFT_OUT will drop back into when a particular pressure is released.

Moving on to the device–logical state mapping with its timed events, we will just consider the case when the minidisk knob is pulled out, as in Fig. 8:

$$\begin{aligned}
LS &= \{1, 2, 3, \ldots\}\\
Ev &= \{\texttt{up}, \texttt{down}\}\\
trigger &= \{<\texttt{CENTRE_OUT}, \texttt{LEFT_OUT}> \mapsto \texttt{down}, <\texttt{CENTRE_OUT}, \texttt{RIGHT_OUT}> \mapsto \texttt{up}\}\\
time\text{-}trigger &= \{<\texttt{LEFT_OUT}, 1sec, \texttt{PERIODIC}> \mapsto \texttt{down}, <\texttt{RIGHT_OUT}, 1sec, \texttt{PERIODIC}> \mapsto \texttt{up}\}\\
doit &= \{<\texttt{up}, 1> \mapsto 2, <\texttt{up}, 2> \mapsto 3, \ldots, <\texttt{up}, 16> \mapsto 17, <\texttt{down}, 17> \mapsto 16, \ldots\}
\end{aligned}$$

7. Controlled state and compliant interaction

7.1. Example: washing machine and electric kettle

Finally we come to devices where the state of the physical device is affected by the underlying logical system as well as vice versa. Consider a washing machine control knob that sets the programme (Fig. 9a) or an electric kettle switch (Fig. 9b). In each case the user can control the device: twisting the knob to set the programme or pushing up or down the kettle switch to turn the kettle on and off. However, in addition the underlying logical system can also control the physical device. In the case of the washing machine as the clothes are washed the dial usually moves round to act as a display of the current state of the programme. In the case of the kettle, when the water boils many kettles both switch themselves off and at the same time release the switch.

We say that this kind of device has *controlled state*; that is the state of the physical device is not simply manipulated as in input by the user, but is also controlled as an output by the underlying logical system.

In fact both systems in addition exhibit *compliant interaction* [GhD03] where the system control of the physical device operates in a compatible way to the user control: with the kettle the user can turn the switch off or the system can and the effect on the switch is the same for both user or system control. Of course there are usually limits to compliant interaction: the kettle does not turn itself on and the user turning the knob to the end of the wash cycle does not magically wash the clothes!

Figure 10 shows the state diagram for the kettle switch and also the state of the power and water. Strictly there are two sub-systems in the kettle: the *power* (ON/OFF) influencing the *water temperature* (continuous

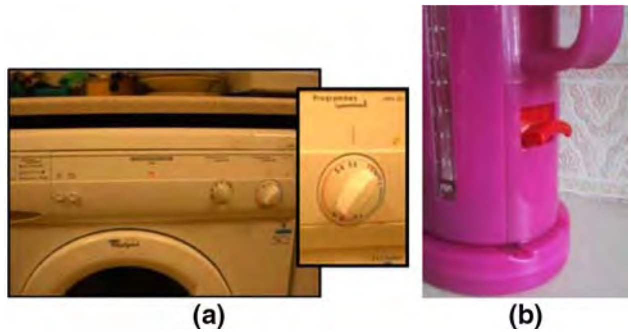


Fig. 9. Compliant interaction: **a** washing machine knob, **b** kettle switch

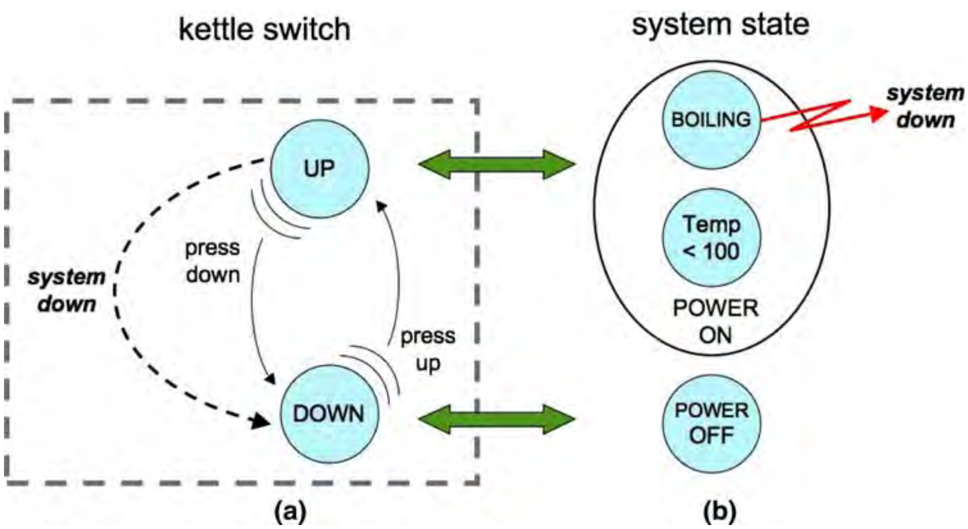


Fig. 10. Electric kettle: **a** kettle switch, **b** power and water

scale), but for simplicity we have shown the water state as simply boiling vs. not boiling and only as sub-states of the POWER-ON state. The arrows between the device and logical state show that there is an exposed state for the electrical power system. The little lightning arrow from the water's BOILING state shows that simply being in the state, by itself, triggers the system action 'system down'. Like user actions in the physical world this is protracted and lasts as long as the kettle is boiling, it is not simply an event at the moment boiling is first sensed. This possibility of an autonomous action is shown by the dashed transition on the state diagram for the physical switch.

Note how the system action and the user action to switch off the kettle are both operating in exactly the same way on the physical device. Note also that if the user is pushing up when the system is trying to switch the kettle off there is a conflict and whether the switch goes off or not depends on who is stronger! For most electric kettles the automatic switching off is usually weaker than the user's ability to hold the switch up (usually simply releasing a catch) so it is possible to boil the kettle when dry. In some kettle designs the power is switched off by the system when the water is boiling irrespective of whether the user allows the switch to go down; in this case we would have similar device states, but different logical state transitions and no exposed state mapping.

Again note that if we used a notation with an in-built model of syncronisation between components, this conflict and alternative designs might be at best missed and at worst mis-specified. This is not to argue that one should not use more elaborate and notations, but that for this investigative analysis the simpler notation forces us to face important design issues.

7.2. Formal model

To deal with these kinds of devices we need to add a set of system actions SA and have a mapping that says which system actions are triggered by which logical states:

$sys\text{-}trigger : LS \rightarrow set(SA)$

These system actions will then have their effect on device state transitions just like user actions:

$sys\text{-}action : SA \leftrightarrow DT$ −n − m partial relation

Just like user actions it is possible that a single system action may have different effects in different device states and that several system actions might be possible in a single device state. However, when it is an exposed state system, like the kettle, it is likely that the system actions are very specific for a particular state. Indeed if there is a state mapping, then there should be some consistency between the system state(s) that correspond to a device state and the system actions pertaining to each:

$$\forall s \in LS \;\; \forall a \in sys\text{-}trigger(s)$$
$$\exists d \in dom(sys\text{-}action(a)) \; st. \; (d, s) \in state\_mapping$$
$$\forall a \in SA \;\; \forall d \in dom(sys\text{-}action(a))$$
$$\exists s \in sys\text{-}trigger^{-1} \; st. \; (d, s) \in state\_mapping$$

Or equivalently:

$$\forall s \in LS \;\; \forall a \in sys\text{-}trigger(s)$$
$$dom(sys\text{-}action(a)) \cap state\_mapping^{-1} \neq \emptyset$$
$$\forall a \in SA \;\; \forall d \in dom(sys\text{-}action(a))$$
$$dom(sys\text{-}trigger^{-1}) \cap state\_mapping \neq \emptyset$$

In each case, the first of these says that if a logical state can trigger a system action then at least one of the device states consistent with that logical state must take account of that system action. The second says the converse, that if a device state can be affected by a system action then it must be possible for one of the logical states consistent with that device state to generate the action.

Either of these conditions may be broken, but that would suggest that some aspect of the physical device is not being fully utilised, or some signal from the logical device is being ignored. This may be an intended effect of the combination, but certainly merits checking.

7.2.1. Modelling the example

The kettle in Fig. 10 can now be modelled:

$$\begin{aligned}
DS &= \{\texttt{UP}, \texttt{DOWN}\} \\
DT &= \{< verb + UP+, \texttt{DOWN} >, < \texttt{DOWN}, verb + UP+ >\} \\
UA &= \{\texttt{PRESS-DOWN}, \texttt{PULL-UP}\} \\
action &= \{\texttt{PRESS-DOWN} \mapsto < \texttt{UP}, \texttt{DOWN} >, \texttt{PULL-UP} \mapsto < \texttt{DOWN}, \texttt{UP} >\} \\
Z &= \{\}
\end{aligned}$$

For this example, the logical system state itself is more complex. There are two sub-systems, power and water, which we represent by abstraction functions:

$$\begin{aligned}
&power : LS \rightarrow PowerState \\
&water : LS \rightarrow WaterState \\
&PowerState = \{\texttt{POWER_OFF}, \texttt{POWER_ON}\} \\
&WaterState = \{\texttt{NOT_BOILING}, \texttt{BOILING}\}
\end{aligned}$$

When, as in this system, the sub-systems are orthogonal (any combination of sub-system states is possible) and between them completely define the logical state, then LS is simply the Cartesian product of the sub-system states ($LS = PowerState \times WaterState$) and the abstraction functions are simply the component mappings.

Given such sub-system mappings we can define what it means for the system to exhibit exposed state relative to a sub-system:

$exposed\text{-}state \; wrt. \; power \equiv (power \circ state\text{-}mapping)$is one-to-one

This would be exactly the case for the kettle if the kettle is one of the simpler kind that allows you to hold the switch down to keep electricity on when the water is already boiling (it is at this point we can model some of the design alternatives):

$$power \circ state\text{-}mapping = \{\texttt{DOWN} \mapsto \texttt{POWER\_OFF}, \texttt{UP} \mapsto \texttt{POWER\_ON}\}$$

Finally we model the system actions:

$$\begin{aligned} SA &= \{\texttt{system-down}\} \\ sys\text{-}trigger &= \{<\texttt{POWER\_ON}, \texttt{BOILING}> \mapsto \{\texttt{system-down}\}\} \\ sys\text{-}action &= \{\texttt{system-down} \mapsto <\texttt{UP}, \texttt{DOWN}>\} \end{aligned}$$

8. Physigrams in use

8.1. Context

Two of the authors are product designers. They are part of the Programme for Advanced Interactive Prototype Research (PAIPR), a research group attempting to create a suite of systems for the development of computer embedded products sympathetic to the designer's mindset and methods. There are a number of other groups working in this area, e.g. Phidgets [GrF01, Phi08], Voodoo Dolls [PSP99], DTools [HKB05], Switcharoo [AvH02], Denim [LaM95]. Unlike the work of these groups PAIPR has a product design focus rather than an electronics or programming base. PAIPR's methods centre around a system of working that involves low-tech keyboard emulation boxes called IE Units wedded to software building blocks [GLH05]. The system allows rapid prototyping without the usually requisite electronics or programming skills. The system has been used to empirically measure the performance of real products against physical and virtual prototypes and this research found that the link between the physical act of holding a product and interaction was more marked than has previously been understood [EvG06] . This has led to the group to become more interested in the precise nature of physicality in the design process hence the work with the physigrams.

The designers were not involved in the development of the work described in previous sections, in particular, they had not previously been exposed to the physical device STNs (physigrams). For the rest of this section we will refer to them as 'the designers' and in contrast describe the authors who were involved in developing the notation as 'the developers'. Both designers and developers are involved in a broader project on understanding physicality in design. So, there is some danger that the designers share more conceptual background with the developers than would be the case with a typical product designer. However for the exploratory purposes of this case study of use, we believe that the fact that the designers were previously not exposed to physigrams is sufficient to ensure valid results.

As part of a project meeting, the designers were first given a short explanation of the concept by the developers (approximately 10–20 min), in particular the developers introduced the notion of studying the physical device 'unplugged' from its digital aspects and some of the diagrammatic examples. The designers were then given an earlier paper [DGR07] that covers largely the same ground as Sects. 2–7. As the designers were not from a computing or mathematical background they were instructed to ignore the parts on the formal specifications, but to read those regarding the formal diagrammatic notations.

Subsequently, and without any further input or aid from the developers, the designers then read the relevant parts of the paper and spent a collaborative session applying the physigrams to an ongoing design project. It should be noted that the designers were not given a brief by the developers, but applied the techniques to a brief and project that they were pursuing for other reasons; that is the physigrams are effectively being used 'in the wild'. Because of this, the developers were not able to tune the brief to exercise all aspects of the physigrams, for example time effects. However, this free use by the designers has the advantage of not being limited to the developers' pre-conceptions; this gave the designers the freedom to explore issues that the developers may not have considered.

The existing brief the designers were working on involved three alternative devices for interacting with the same underlying application; this was in fact a good exercise for the physigrams as we were able to see how superficially similar, but subtly different devices were distinguishable in the physigrams. The designers were not given any time limit for using the physigrams, but in the end spent around an hour in total during which time they produced initial handsketched physigrams followed by two electronic variants of each of three design alternatives. Of this hour about half the time was spent in discussion and the other half producing the final physigrams.

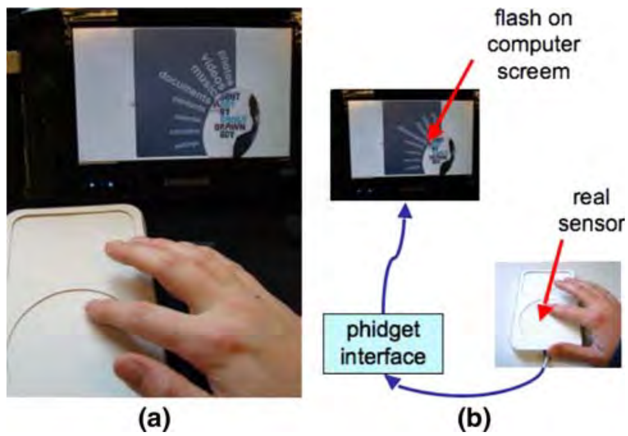


Fig. 11. Prototyping device

Fig. 12. Variants of prototyping device: **a** device 1, **b** device 2 and 3, **c** device 4

8.2. Prototypes

The design exercise they chose was one connected to the shared project involving creating multiple high-fidelity prototypes of an iPod-like device. Figure 11a shows one of the devices in action. The device has a round touch-sensitive area, rather like a circular trackpad, and also above it a rectangular area where the display would be. For the prototype the display is instead emulated in Flash on a separate computer screen, which shows the part-circular menu envisaged for the device.

The prototype is interactive as there is a real touch sensor inside the cardboard mock-up. The sensor is a Phidget [GrF01, Phi08] and the hardware and libraries supplied convert the raw sensor inputs into Flash events. Figure 11b shows the setup.

In all there are four versions of the prototype device, all with identical display functionality, differing only in the input device. Figure 12 shows the devices. Device 1 has a clicking rotary switch with a knob connected to an IE unit [GLH05] that can be turned to 12 different directions. Device 2 has a clicking rotary switch identical to device 1 connected to an IE unit, but it is turned using a flat dial rather than a knob. Note that the knob has a direction that can be detected visually and by touch. In contrast device 2 does not have any distinguishable direction. Device 3 is identical in appearance to device 2 (hence a single photo), but inside has a different rotary switch that rotates freely, but detects 12 orientations, again attached to an IE unit. Finally, device 4 is the one introduced in Fig. 11 with a touchpad.

All the devices also can be pressed to act as a 'select' event. In devices 1, 2 and 3 the whole knob or dial is depressed. In device 4 this is achieved by touching in the centre (NB. this had not been implemented in the prototype but for the purpose of the physigrams the design intent was to have the touch select in the middle).

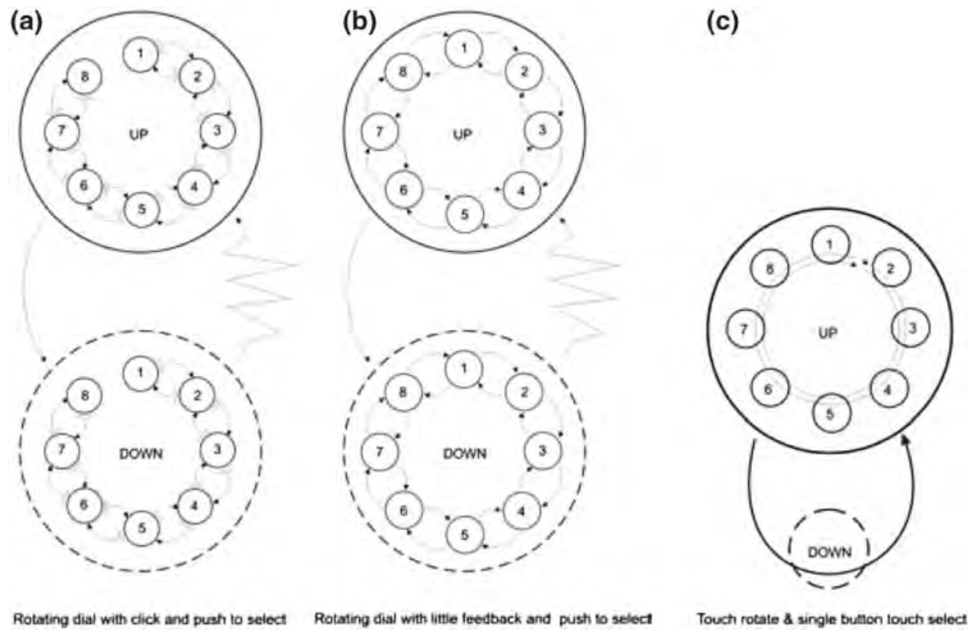


Fig. 13. Physigrams for each device: **a** device 1 and 2, **b** device 3, **c** device 4

8.3. Physigrams

Figure 13 shows the physigrams produced by the designers for each device. Devices 1 and 2 look different and have a visually different control, but have the same rotary switch connected to an IE unit inside and hence the same physigram (Fig. 13a). The physigram shows 16 states in total, 8 when the device is 'up' and 8 when it is 'down'. As the press down on the device is sprung there is a bounce back between the down states and the up states. The designers took some liberty with the notation here and used the circle drawn round the 8 up states to bind them into a single 'super state'. The implied semantics is that when you press up or down you end up in the same numbered state.

Note that there are only 8 states drawn yet there are 12 outputs of the IE unit. When discussing these, at first the developers thought this was a mistake and the designers had not fully grasped the 'unplugged' concept. If you turned the device 12 steps could clearly be felt, the 8 seemed to refer to logical states of the application. Indeed, the designers explained that initially they had drawn the diagram with 8 states labelled by the system functions they selected. However, they then realised that this was not an unplugged device. They considered drawing in 12 states each in the up and down circles, but in the end decided not to. This would have been an accurate description of the rotary switch used in the *prototype*—however, this was just a prototype and the intention was that if the device were actually produced an 8 state dial would have been used. That is they used the physigrams to specify the intended design not the limitations of the prototype.

Device 3 (Fig. 13b) is quite similar in broad terms. The most obvious difference is that there is no state 1 to state 8 transition for devices 1 and 2 as the dials do not rotate a full 360 degrees, whereas device 3 can rotate totally freely. Device 4 differs more radically still. Both device 3 and device 4 have free movement, however the difference is that whilst with device 3 (the freely tuning dial) it is possible to turn the dial whilst it is pressed in, in the case of device 4 (the touchpad) the press has to be in the middle, so there is no rotary movement in the 'down' state.

In fact, the physigram for device 3 also differs from devices 1 and 2 at the level of movement between the sub-states. This is evident in the close-up details in Fig. 14. Devices 1 and 2 have definite 'stops'. As you try to turn the knob or dial there is slight resistance and then when you twist sufficiently the knob/dials clicks to the next position. That is, it is a bounce-back/give behaviour and the designers used the shorthand introduced in Sect. 5.3.

Looking at the detail for device 4 (Fig. 14c) recall that the user's finger can move freely over the trackpad. The designers indicated this by the continuous circles, showing the finger can move freely in either direction. However,

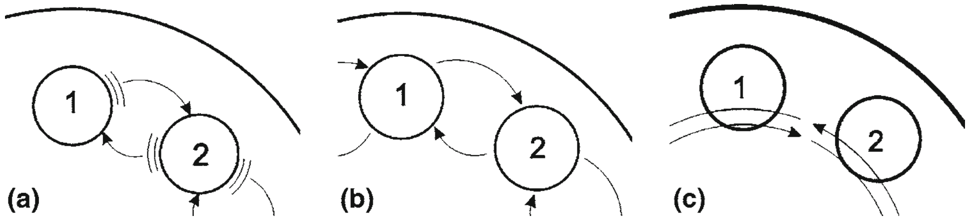


Fig. 14. Detail from physigrams for each device: **a** detail device 1 and 2, **b** detail device 3, **c** detail device 4

some locations are sensed by the system and used to represent states. The designers drew numbered states on this overlaying the continuous circles to show that different areas had different meanings. Here they again seemed to be not taking on board the 'unplugged' concept. However, they explained that they had considered this carefully and had chosen not to label the states with application names (as they had decided for devices 1 and 2), but they also knew that in the hardware the device reported only 8 states and so they used the diagram to represent this. This highlighted an ambiguity in unplugged-ness. There are actually three levels:

1. the physical device fully connected with its underlying application
2. the device with its internal electronics, but unplugged from its application
3. the purely physical aspects of the device

The physigrams were originally intended for level 1, but the designers had used them for level 2. Both are of course important. Indeed, this is exactly the same kind of issues that we found when considering timed transitions in Sect. 6. There the τ annotation in Fig. 8a referred precisely to level 2. This may mean that for certain devices we should consider drawing both level 2 and level 3 diagrams or perhaps annotating a single diagram to make clear which elements are level 2.

A problem with the use of level 2 physigrams, which we noted in Sect. 6, is that they embody knowledge that is available to the designers, but may not be apparent to users. For the timed transitions this was simply annotations to states, but here there are effectively states drawn that are not part of the physical device behaviour. However, insisting on level 1 diagrams would be counter productive as clearly describing level 2 features are of value also. Many applications used by practicing designers use notions of layers (e.g. Photoshop); so, one could envisage tool support that allowed multiple layers of annotations to be added to a basic physigram thus encouraging designers to consider both the raw device behaviour and also the way this interacts with low-level digital features.

From the description it would appear that device 3 and device 4 should have similar physigrams for the rotation part as both have totally free movement to any orientation. In fact the detailed view of device 3 looks more similar to devices 1 and 2. The designers explained that although the dial did not have click stops in the way that device 2 did, in fact it was just possible to feel when the dial moved past one of the contacts. That is while device 2 had haptic feedback (actual resistance) device 3 has tactile feedback (felt transitions). The designers had represented this by using a simple state change diagram for device 3 compared to the bounce-back in device 2 (similar state transitions, but different arrow shape). This makes a clear distinction between the two, but is not entirely satisfactory as it does not show the continuity of movement possible in the way that the physigrams for device 4 does. This is another example, as we have found previously in this paper, of the importance of being able to deal adequately with continuous phenomena.

8.4. Using the page

The formal meaning of the physigrams depends only on the topology and connectivity of states and arrows and not their precise positions. This meant that the designers were free to use spatial layout on the page to convey additional meaning. In some cases this represented aspects that one might want to capture formally, in particular the idea that the 'big' up/down transitions in Fig. 13a, b take you to the corresponding numbered sub-states. More often the page is used to help the human reader make sense of the diagram, making the form on the paper correspond roughly to the form of the device, as in the circular dial.

The designers in fact went through several iterations before ending up with the neat versions in Fig. 13, some of which was about interpretation of the more formal or semantic issues (such as whether to label the states abstractly 1–8 or by application labels), but much was also about the most useful layout.

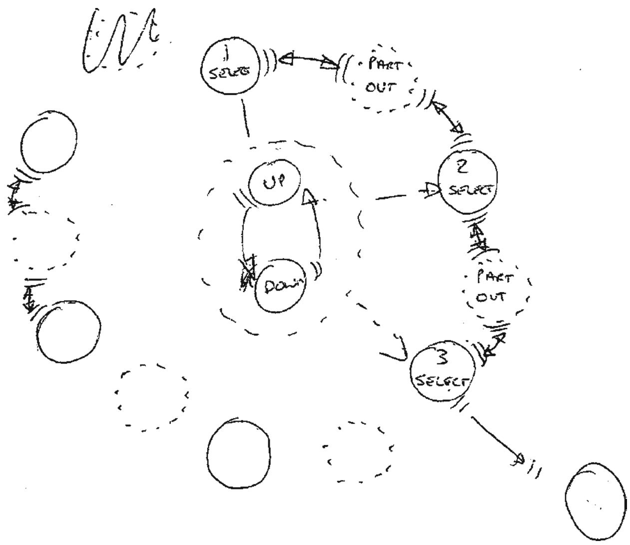


Fig. 15. Early physigrams sketching

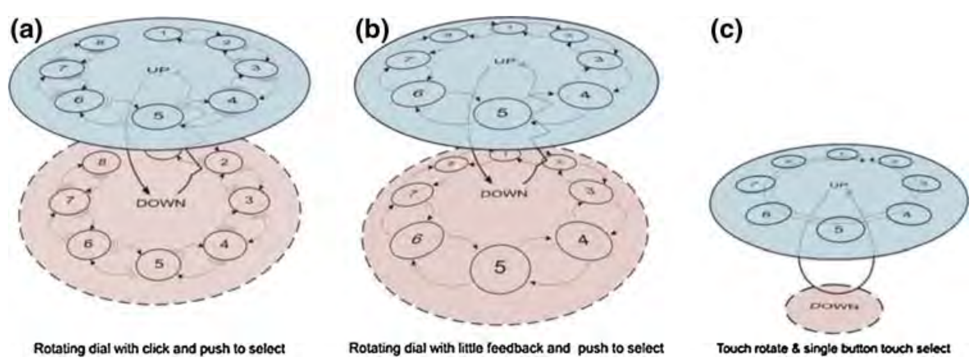


Fig. 16. Physigrams with perspective: **a** device 1 and 2, **b** device 3, **c** device 4

Figure 15 shows an early sketch for device 1. We can see examples of both semantic decisions and human representation ones. Looking round the outside it is clear that the designers were considering whether to represent explicitly intermediate states between the click stops, (labelled ‘part out’), a clear semantic decision. However, also at this stage we see the up/down transition being represented in the middle (rather like a concurrent state diagram in state charts).

This use of space was most complete in a set of alternative 3D perspective physigrams (Fig. 16) that the designers produced. Here they explicitly explained how the layering of the up and down super states conveyed the fact that down was literally pressing down.

It is tempting in a visual notation to try to use the 2D layout to convey explicit semantics, for example, some notations use juxtaposition to mean communication, or left-to-right layout as sequence. Clearly some level of this is useful conveying formal semantics implicitly using the human visual system. However, it is the very fact that the formal semantics of the physigrams left much of the 2D layout *uninterpreted* that makes it available for additional human interpretation. This parallels in notation design, one of the general principles of designing for appropriation: ‘allow interpretation’ [Dix07]. The importance of such ‘secondary notation’ (annotation or variants in notations that do not impact on semantics) has also been recognised in the cognitive dimensions literature providing, *inter alia*, means to add more nuanced human-interpretable aspects to an underlying formal notation [GrP96].

It was perhaps a little surprising that none of the physigrams included photographs or sketches of the visible aspects of the devices. The diagrams did explicitly encode the haptic and tactile differences between devices 2 and 3 and between device 3 and 4, but not the visual difference between device 1 and 2. Device 1 and 2 share an identical physigram, but device 1 is an exposed state device whereas device 2 is a hidden state device. This does not show up in the state diagrams for the physical device, but would do in the mapping between device states and visible states. During discussion it became clear that the physigrams had not been annotated with the visible form (as in Fig. 2a) because the physical devices were there in front of the designers. The visual aspects were immediately obvious and did not need to be formally recorded, however the aspects that were less obvious were more worthwhile to express in the physigram. This was because the physigrams were being used as a live tool during a design meeting. If they were used to communicate between different teams—as when the developers tried to interpret them—it would be more important to give explicit guidance on visual annotations.

8.5. Designers' impressions

It is important to note that generally designers do not normally work this way, it is outside their comfort zone. That said, they were able to work with the concept quite quickly, producing the first finished physigram within about 20 minutes and the new 3D physigram concept within an hour of starting the task.

When initially asked, the design team envisaged using physigrams as a retrospective descriptive tool, probably because this was effectively how they were being used in the exercise. However, after further consideration, the designers suggested that there would be more value in deploying physigrams to describe the interaction early in the design process when the real interaction cannot be prototyped. This would aid communication within the design team and perhaps help individual designers' thought processes, describing and analysing how an interaction should be before the prototyping stage.

The designers also speculated on how it would be to apply this technique to a whole product rather than just a simple dial. They wondered how extra interactions would be conveyed, as separate diagrams or a complete diagram for the product? They also questioned whether interactions would be represented completely separately and whether this actually helps communication between design teams or introduces complications. Effectively they were asking questions about the way a notation could handle both details of the physigrams and also the complexity of the system as whole; questions familiar to those involved in many kinds of formalism.

Considering physigrams as a communication tool for the design process, the designers appeared to feel most comfortable working at level 2, described previously as 'the device with its internal electronics, but unplugged from its application', as evidenced by their representation of device 4. This may have been influenced by the fact that the prototypes used the IE system to link to a Flash animation and were thus very aware of the events generated by the different IE units. This level of analysis helped them make distinctions between the device as prototyped (with all the limitations of the off-the-shelf components) and the device as envisaged in production, so was valuable in that respect. However, as discussed earlier, the disadvantage of this level of analysis is that they were effectively encoding information in the physigram that would not be apparent to a user simply picking up the physical device. The level 1 description would have potentially sensitised the designers to these 'pick up and use' aspects of the device. More work is clearly needed to establish the best form and level of this kind of specification, but the proposed use of layers may be a solution.

9. Discussion

We have used a number of examples to show different ways in which the physical states of a device can interact with the logical states of the system. These have reinforced the importance of distinguishing the two and being able to talk about the, sometimes quite subtle, differences between what might appear to be similar controls.

Each example has dealt with a different property that we have introduced in previous work: exposed state, hidden state, bounce-back, controlled state and compliant interaction. For each of these we have (i) discussed examples informally, then (ii) expressed the examples using parallel state transition networks for the physical and logical states and (iii) given the STNs and their relationship semantics using a formal model. We have introduced the formal model piecewise as each property requires additional elements in the model.

For practical design the variants of STNs would seem more appropriate than the model, although the latter gives the former a more precise semantics. The simpler examples and the relationship between the physical and logical STNs could be dealt with using standard notations and certainly could be translated into state-charts or similar notations. However, as we looked at more complex properties such as bounce-back we had to extend standard state-transition networks to represent the additional effects. This exposes design issues such as the appropriate use of tension states.

When real designers used the physigrams we found that even the two level distinction between physical 'unplugged' device and the logical states was not sufficient. We found we also needed to consider the device at an intermediate level 'unplugged' and yet with its internal digital aspects. While we had already had intimations of this in time-dependent events, the additional complexity of the novel prototypes exposed this and additional issues, highlighting the importance of exposing formal work to empirical study.

Possibly most surprising for the designers was the developers' use of flexibility in the layout of the formal notation to add additional informal interpretations. In retrospect this was fully in accordance with design guidance for appropriation of artifacts. As a general rule this suggests that formal notations should attempt to leave aspects without a formal interpretation, in particular layout; thus leaving these aspects open to human interpretation.

Issues of continuity have arisen repeatedly throughout this paper, both in the standard device examples and in the design case study. Human action and physical device interaction is not simply a matter of 'events' occurring and their discrete effects on state. In real life we interact continuously and experience continuous responses; we exert force and feel pressure. However, we also experience discontinuous effects, both with physical devices (when the light switch snaps to a new position) and even more so in digital interactions. This suggests that a deeper semantics based on status–event analysis is still needed in order to map the still discrete formal modelling of this paper into something approaching the physics of real life.

A specific example of continuity in physical interaction was the importance of 'give' in several devices. While mentioned obliquely in the literature, this does not appear to have yet had the attention it deserves in allowing exploration of the action potential of devices. Purely digital interactions rarely have this ability, however there are some interaction techniques that have some of this quality. For example in the most recent version of Microsoft Office transparent context-dependent tool palettes become more opaque when the mouse moves towards them.

The haptic feedback of this 'give', along with the (only just) perceptible tactile feedback of some devices, and the way physical constraints become a form of dialogue, all point the way to effective use of physicality to guide or constrain digital interactions.

Returning to our two goals. We have obtained new insights and understanding based on the analysis, for example, the issue of 'give' above and the way this cast new light on sequential affordances. We have also seen that physigrams have potential within the design process although this requires further study and tool support.

We chose to restrict our scope to the behavioural aspects of the physical device. However, various other physical aspects of the device are important. These include CAD diagrams, which are used extensively in product design; detailed layout [Thi07], the human control loop [E06] and the disposition of the device in its environment [RDR05]. Looking forward it would be valuable to seek ways to link these different aspects. A single notation for all is likely to be cumbersome, but a patchwork of notations and representations could be linked by a unifying abstract semantics. The development of such a unifying semantics would be a major challenge involving representations of continuous action, physical pressure and the physics of the environment itself.

Acknowledgments

This work has been supported by the AHRC/EPSRC funded project DEPtH 'Designing for Physicality' (http://www.physicality.org/), part of the joint AHRC/EPSRC research initiative 'Designing for the 21st Century' (http://www.design21.dundee.ac.uk/).

References

[Ans92] Anson E (1992) The device model of interaction. SIGGRAPH Comput Graph 16(3):107–114

[AvH02] Avrahami D, Hudson S (2002) Forming interactivity: a tool for rapid prototyping of physical interactive products. In: Proc. of the 4th Conf. on Designing interactive Systems: Processes, Practices, Methods, and Techniques (DIS '02). ACM, New York, pp 141–146. http://doi.acm.org/10.1145/778712.778735

[BSK03] Benford S, Schnadelbach H, Koleva B, Gaver B, Schmidt A, Boucher A, Steed A, Anastasi R, Greenhalgh C, Rodden T, Gellersen H (2003) Sensible, sensable and desirable: a framework for designing physical interfaces, Technical Report Equator-03-003, Equator, 2003. http://www.equator.ac.uk/

[BoV90] Booker S, Vertelney L (1990) Designing the whole-product user interface. In: Laural B (ed) The art of computer interface design. Addison-Wesley, Reading, pp 57–63

[Bux86] Buxton W (1986) There's more to interaction than meets the eye: some issues in manual input, In: Norman D, Draper S (eds) User centered system design: new perspectives on human–computer interaction. Lawrence Erlbaum Associates, Hillsdale, pp 319–337, [Bux90] Buxton W (1990) A three-state model of graphical input. In: Proc. of human–computer interaction—INTERACT '90. Elsevier, Amsterdam, pp 449–456

[CMR90] Card S, Mackinlay J, Robertson G (1990) The design space of input devices. In: Proc. of CHI'90. ACM Press, New York, pp 117–124

[CMR91] Card S, Mackinlay J, Robertson G (1991) A morphological analysis of the design space of input devices. ACM Trans Inf Syst 9(2):99–122

[Car94] Carr D (1994) Specification of interface interaction objects. In: Proc. of CHI '94. ACM, New York, pp 372–378

[CoN06] Coutrix C, Nigay L (2006) Mixed reality: a model of mixed interaction. In: Proc. of AVI'06. ACM Press, New York, pp 43–50

[CoN08] Coutrix C, Nigay L (2008) Balancing physical and digital properties in mixed objects. In: Proc. of AVI'08, the Working Conf. on Advanced Visual interfaces. ACM Press, New York, pp 305–308

[CuR07] Curzon P, Rukšėnas R, Blandford A (2007) An approach to formal verification of human-computer interaction. Formal Aspects Comput 19(4):513–550

[Deg04] Degani A (2004) Taming HAL: designing interfaces beyond 2001. Palgrave Macmillan, New York

[DiR85] Dix A, Runciman C (1985) Abstract models of interactive systems. People and computers: designing the interface. Cambridge University Press, Cambridge, pp 13–22

[Dix91] Dix A (1991) Formal methods for interactive systems. Academic Press, New York. http://www.hiraeth.com/books/formal/

[Dix91b] Dix A (1991) Status and events: static and dynamic properties of interactive systems. In: Proc. of the Eurographics Seminar: Formal Methods in Computer Graphics. http://www.hcibook.com/alan/papers/euro91/

[DiA96] Dix A, Abowd G (1996) Modelling status and event behaviour of interactive systems. Softw Eng J 11(6):334–346 (1996). http://www.hcibook.com/alan/papers/SEJ96-s+e/

[DiA96b] Dix A, Abowd G (1996) Delays and temporal incoherence due to the mediated status–status mappings. SIGCHI Bull 28(2):47–49

[Dix03] Dix A (2003) Getting physical, keynote at: OZCHI 2003, Brisbane, Australia. http://www.hcibook.com/alan/talks/ozchi2003-keynote/

[DFA04] Dix A, Finlay J, Abowd G, Beale R (2004) Human–computer interaction, 3rd edn. Prentice Hall, Englewood Cliffs. http://www.hcibook.com/e3/

[Dix07] Dix A (2007) Designing for appropriation. In: Procedings of BCS HCI 2007, People and Computers XXI, vol 2, BCS eWiC. http://www.bcs.org/server.php?show=ConWebDoc.13347

[DGR07] Dix A, Ghazali M, Ramduny-Ellis D (2007) Modelling devices for natural interaction. In: Proc. of Second Intnl. Workshop on Formal Methods for Interactive Systems, FMIS2007, ENTCS. Elsevier, Amsterdam

[DLF07] Dix A, Leite J, Friday A (2008) XSED—XML-based description of status–event components and systems. In: Proc. of Engineering Interactive Systems 2007 (EIS 2007). Lecture notes in computer science, vol 4940. Springer, Berlin

[DSG02] Dubois E, Silva P, Gray P (2002) Notational support for the design of augmented reality systems. In: Proc. of the 9th International Workshop on interactive Systems. Design, Specification, and Verification, DSVIS2002. Lecture notes in computer science, vol 2545. Springer, Berlin, pp 74–88

[DuG07a] Dubois E, Gray P (2008) A design-oriented information-flow refinement of the ASUR interaction model. In: Engineering interactive systems (incorporating EHCI, HCSE, DSV-IS). Lecture notes in computer science, vol 4940. Springer, Berlin

[DGR07b] Dubois E, Gray P, Ramsay A (2007) A model-based approach to describing and reasoning about the physicality of interaction. In: Proc. of Physicality 2007. UWIC Press, Cardiff, pp 77–82

[E06] Eslambolchilar P (2006) Making sense of interaction using a model-based approach. Ph D thesis, Hamilton Institute, National University of Ireland, NUIM, Ireland

[EvG06] Evans M, Gill S (2006) Rapid development of information appliances. In: Proc. of International Design Conf. Design 2006, (Croatia, 15–18 May 2006)

[Gav91] Gaver W (1991) Technology affordances. In: Proc. of CHI '91. ACM Press, New York, pp 79–84

[GhD03] Ghazali M, Dix A (2003) Aladdin's lamp: understanding new from old. In: Proc. of 1st UK-UbiNet Workshop, Imperial College London. http://www.hcibook.com/alan/papers/ubinet-2003/

[GhD05] Ghazali M, Dix A (2005) Visceral interaction. In: Proc. of the 10th British HCI Conf., vol 2, pp 68–72. http://www.hcibook.com/alan/papers/visceral-2005/

[GhD06] Ghazali M, Dix A (2006) Natural inverse: physicality, interaction & meaning. In: Let's Get Physical: Tangible Interaction and Rapid Prototyping in, for, and about Design Workshop at 2nd International Conf. on Design Computing & Cognition 2006

[Gib86] Gibson J (1986) The ecological approach to visual perception. Houghton Mifflin, USA

[GLH05] Gill S, Loudon G, Hewett B, Barham G (2005) How to design and prototype an information appliance in 24 hours—integrating product & interface design processes. In: Proc. of the 6th International Conf. on Computer Aided Industrial Design and Concept Design, University of Delft, The Netherlands

[GrP96] Green T, Petri M (1996) Usability analysis of visual programming environments: a 'cognitive dimensions' framework. J Vis Languages Comput 7:131–174

[GrF01] Greenberg S, Fitchett C (2001) Phidgets: easy development of physical interfaces through physical widgets. In: Proc. of the 14th Annual ACM Symposium on User interface Software and Technology (UIST '01), pp 209–218. http://doi.acm.org/10.1145/502348.502388

[Har87] Harel D (1987) Statecharts: a visual formalism for complex systems. Sci Comput Program 8(3):231–274

[HKC07] Harrison M, Kray C, Campos J (2007/2008) Exploring an option space to engineer a ubiquitous computing system. In: 2nd International Workshop on Formal Methods for Interactive Systems (FMIS 2007), Electronic Notes in Theorectical Computer Science, vol 208. Elsevier, Amsterdam, pp 41–55

[HKB05] Hartman B, Klemmer S, Bernstein M, Mehta N (2005) d.tools: Visually Prototyping Physical UIs through Statecharts. In: Extended Abstracts of UIST 2005. ACM, New York

[Har03] Hartson H (2003) Cognitive, physical, sensory, and functional affordances in interaction design. Behav Inform Technol 22(5):315–338

[Ish08] Ishii H (2008) Tangible user interfaces. In: Sears A, Jacko J (eds) The human–computer interaction handbook fundamentals, evolving technologies, and emerging applications, Chapter 24, 2nd edn. Laurence Earlbaum, London, pp 469–487
[JDM99] Jacob J, Deligiannidis L, Morrison S (1999) A software model and specification language for non-WIMP user interfaces. ACM Trans Comput Hum Interact 6(1):1–46
[Joh96] Johnson C (1996) The evaluation of user interface design notations. In: Proc. of Design, Specification and Verification of Interactive Systems '96. Springer, Berlin, pp 188–206. http://www.dcs.gla.ac.uk/~johnson/papers/chris_jarle/
[LaM95] Landay J, Myers B (1995) Interactive sketching for the early stages of user interface design. In: Proc. of CHI'95. ACM Press/Addison-Wesley, New York, pp 43–50. http://doi.acm.org/10.1145/223904.223910
[LoH02] Loer K, Harrison M (2002) Towards usable and relevant model checking techniques for the analysis of dependable interactive systems. In: Proc. 17th International Conf. on Automated Software Engineering. IEEE Computer Society, New York, pp 223–226
[MDS99] Massink M, Duke D, Smith S (1999) Towards hybrid interface specification for virtual environments. In: DSV-IS 1999 Design, Specification and Verification of Interactive Systems. Springer, Berlin, pp 30–51
[Mil88] Milner N (1988) A review of human performance and preferences with different input devices to computer systems. In: Proc. of HCI88, People and Computers IV. Cambridge University Press, Cambridge, pp 341–362
[MDB96] Moher T, Dirda V, Bastide R, Palanque P (1996) Monolingual, articulated modelling of users, devices and interfaces. In: 3rd EUROGRAPHICS workshop on design, specification and verification of Interactive systems. Springer, Berlin, pp 312–329
[Mon86] Monk A (1986) Mode errors: a user-centered analysis and some preventative measures using keying-contingent sound. Int J Man Mach Stud 24(4):313–327
[NiM94] Nielsen J, Mack R (1994) Usability inspection methods. Wiley, New York
[NCo91] Nigay L, Coutaz J (1991) Building user interfaces: organizing software agents. In: ESPRIT '91 Conf., pp 707–719
[NCo95] Nigay L, Coutaz J (1995) A generic platform for addressing the multimodal challenge. In: Proc. of CHI'95. ACM, New York, pp 98–105
[Nor99] Norman D (1999) Affordance, conventions, and design. Interactions 6(3):38–43
[PaP97] Palanque P, Paterno F (1997) (eds) Formal methods in human–computer interaction. Springer, Berlin
[Pa69] Parnas D (1969) On the use of transition diagrams in the design of a user interface for an interactive computer system. In: Proc. of the 1969 24th National Conf.. ACM, New York, pp 379–385
[PaG86] Payne S, Green T (1986) Task–action grammars: a model of mental representation of task languages. Hum Comput Interact 2(2):93–133
[PfH85] Pfaff G, Hagen P (1985) (eds): Seeheim workshop on user interface management systems. Springer, Berlin
[Phi08] Phidgets Inc., 2008. http://www.phidgets.com/
[PSP99] Pierce J, Stearns B, Pausch R (1999) Voodoo dolls: seamless interaction at multiple scales in virtual environments. In: Proc. of the 1999 Symposium on Interactive 3D Graphics, pp 141–145
[RDR05] Ramduny-Ellis D, Dix A, Rayson P, Onditi V, Sommerville I, Ransom J (2005) Artefacts as designed, Artefacts as used: resources for uncovering activity dynamics. Cogn Technol Work 7(2):76–87
[Rei81] Reisner P (1981) Formal grammar and human factors design of an interactive graphics system. IEEE Trans Softw Eng SE-7(2):229–240
[Shn83] Shneiderman B (1983) Direct manipulation: a step beyond programming languages. IEEE Comput 16(8):57–69
[Shn84] Shneiderman B (1984) Response time and display rate in human performance with computers. ACM Comput Surv 16(3):265–285
[Smi06] Smith S (2007) Exploring the specification of haptic interaction. In: Interactive systems: design, specification and verification (DSVIS 2006). Lecture notes in computer science, vol 4323. Springer, Berlin, pp 171–184
[Suf82] Sufrin B (1982) Formal specification of a display editor. Sci Comput Program 1:157–202
[ThH90] Thimbleby H, Harrison M (1990) Formal methods in human–computer interaction. Cambridge University Press, Cambridge
[Thi07] Thimbleby H (2007) Press On: principles of interaction programming. MIT Press, Cambridge
[Thi07] Thimbleby H (2007) Using the Fitts law with state transition systems to find optimal task timings. In: Pre-Proc. of Second Intnl. Workshop on Formal Methods for Interactive Systems, FMIS2007. http://www.dcs.qmul.ac.uk/research/imc/hum/fmis2007/preproceedings/FMIS2007preproceedings.pdf
[UIM92] UIMS (1992) A metamodel for the runtime architecture of an interactive system: the UIMS tool developers workshop. SIGCHI Bull 24(1):32–37
[UIJ05] Ullmer B, Ishii H, Jacob R (2005) Token+constraint systems for tangible interaction with digital information. ACM Trans Comput Hum Interact 12(1):81–118
[WDO04] Wensveen S, Djajadiningrat J, Overbeeke C (2004) Interaction frogger: a design framework to couple action and function. In: Proc. of the DIS'04. ACM, New York, pp 177–184
[WiH00] Willans J, Harrison M (2001) Verifying the behaviour of virtual world objects. In: Palanque P, Paternó F (eds) Proc. of DSV-IS'2000. Springer, Berlin, pp 65–77
[Wüt99] Wüthrich C (1999) An analysis and model of 3D interaction methods and devices for virtual reality. In: Proc. of DSV-IS'99. Springer, Berlin, pp 18–29
[YGS89] Young R, Green T, Simon T (1989) Programmable user models for predictive evaluation of interface design. In: Proc. of CHI'89: Human Factors in Computing Systems. ACM Press, New York

Received 13 March 2008
Accepted in revised form 30 October 2008 by A. Cerone, P. Curzon and D.A. Duce
Published online 6 December 2008

Exploring Physicality in the Design Process

Devina Ramduny-Ellis, InfoLab21, Lancaster University, United Kingdom

Joanna Hare, PDR, University of Wales Institute, Cardiff, United Kingdom

Alan Dix, InfoLab21, Lancaster University, United Kingdom

Steve Gill, PDR, University of Wales Institute, Cardiff, United Kingdom

Abstract

The design process used in the development of many products we use daily and the nature of the products themselves are becoming increasingly digital. Although our whole world is turning ever more digital, our bodies and minds are naturally conceived to interact with the physical. Very often, in the design of user-targeted information appliances, the physical and digital processes are formulated separately and usually, due to cost factors, they are only brought together for user testing at the end of the development process. This not only makes major design changes more difficult but it can also significantly affect the users' level of acceptance of the product and their experience of use. It is therefore imperative that designers explore the relationship between the physical and the digital form early on in the development process, when one can rapidly work through different sets of ideas. The key to gaining crucial design information from products lies in the construction of meaningful prototypes. This paper specifically examines how physical materials are used during the early design stage and seeks to explore whether the inherent physical properties of these artefacts and the way that designers interpret and manipulate them have a significant impact on the design process. We present the findings of a case study based on information gathered during a design exercise. Detailed analysis of the recordings reveals far more subtle patterns of behaviour than expected. These include the ways in which groups move between abstract and concrete discussions, the way groups comply with or resist the materials they are given, and the complex interactions between the physicality of materials and the group dynamics. This understanding is contributing to ongoing research in the context of our wider agenda of explicating the fundamental role of physicality in the design of hybrid physical and digital artefacts.

Keywords

Physicality; Digitality; Product Design; Design Process; Prototyping; Materials

Traditional product design focused on purely physical artefacts designed using physical materials such as clay, wood or plastic foam. As humans we are fitted to live in a physical environment and the behaviour of stone and wood, water and metal appear 'natural' to us either through genetic make-up or early development. However, this is changing as the artefacts designed increasingly include digital elements, from MP3 players to electric drills, and in the design process itself, pencil sketches and clay models often give way to CAD.

It is in this context that we are seeking to explicate the properties of physical materials and physical artefacts and the way we understand and manipulate them, so that we can (i) better inform the design of hybrid digital/physical artefacts and (ii) understand the impact of changing tools and techniques on the design process.

Our previous work has considered properties of commercially available electronic and domestic products in order to uncover the ways in which designers have (maybe tacitly) exploited the physical nature and placement of controls such as knobs and buttons. Often quite subtle differences have a major influence on the naturalness of interaction for end-users (Ghazali & Dix, 2005), that is the extent to which the product exploits the user's automatic or subconscious reactions and behaviours. Another study of mobile phone prototypes showed that physical mock-ups of the interfaces can generate significantly more useful user feedback compared with purely on-screen interactive prototypes; whilst this was as expected, more surprisingly, even quite crude physical mock-ups were as useful as high fidelity ones (Gill et al., 2005).

In this paper, our focus is on physicality within the design process itself. We present the findings of a case study, based on information gathered during a design exercise we ran at the Second International Workshop on Physicality (Ramduny-Ellis et al., 2007) held at Lancaster University as part of the HCI 2007 conference. The objective of the exercise was to explore the role that physicality plays during the design process by setting a common design brief, and providing groups with different materials with which to solve it. Through doing this we hope to better understand the way tacit or explicit properties of these physical materials affect the process and outcomes of early design exploration.

The next section presents the motivation behind our work on understanding the fundamental role of physicality in product design. We then describe the case study in detail and walk through the design process of one of the teams. A detailed analysis of the video recordings of each team design activity revealed several recurrent themes and issues, which enabled us to unpick the rich interplay between materials, design brief, team makeup and dynamics. This thematic discussion is addressed in the final section.

Motivation

Human centred development of computer embedded products, and more specifically user targeted 'information appliances' (products embedded with computers such as mobile phones, digital cameras) are at the crossroads of a number of disciplines (Norman, 1998); therefore their development process can frequently be disjointed with the physical and digital interactions being designed in isolation. These physical and digital processes are often only combined for user testing near the end of the development process when major design changes are impossible. Baxter (2002), and Branham (2000) identified the need for new tools to overcome the problem.

In order to create an effective and a pleasurable experience for the user, designers need to ensure that the relationship between the physical and digital form is explored thoroughly at the early stages of the design process when ideas can be worked through quickly. Methods such as Experience Prototyping (Buchenau & Suri, 2000) and Paper prototyping (Snyder, 2003) go someway towards answering this issue, particularly in their inclusion of physicality; however, methods that retain their "quick and dirty" hands-on approach while incorporating more accurate simulation are still required. One of the problems facing designers in modern businesses is that the pressure to deliver to very tight deadlines and within tight profit margins means that physical prototypes are frequently not produced or are produced with limited functionality or at a low fidelity level. The ability to build meaningful prototypes without investing large sums of money and time is key to obtaining significant design information from product and user testing.

Increasingly, the products we use are a synthesis of digital and physical elements and, for the user, these become indistinguishable. As hybrid physical/digital products are developed, designers have to understand what is lost or confused by this added digitality – and so need to understand physicality more clearly than before.

Our concern with the nature of the physical world and our interaction with it is not new; it has been a concern for philosophers for many years, most notably Heidegger, and is the topic of ongoing debate, particularly related to issues of the embodied mind (Clark, 1998; Wheeler, 2005). For some within psychology, the traditional 'inside-to-outside' Cartesian conceptions of cognition have given way to an increasing acceptance of the importance of physical embodiment for cognition. This is explicit in frameworks such as distributed cognition (Hutchins, 1995), where the role of physical artefacts and multiple actors is seen as essential for 'cognition' to occur and also in the concept of situated action in Suchman's early work at Xerox (Suchman, 1987), which was seminal in bringing ethnographic approaches into interaction research and practice. Environmental psychologists such as Gibson have also explored this area and Gibson's concept of affordance, the ways in which an object is fitted for human action, has entered the vocabulary of interaction design (Gibson, 1979).

Whilst some of these have been applied in design settings, we feel that the range of properties covered does not fully encompass all that is important when using physical materials. For example, the temporal continuity of physical items is taken for granted. Furthermore, the focus is largely on the use of products, however the creative act of design involves a combination of imaginative and manipulative processes. To what extent is this creativity enhanced or inhibited by the physical nature of design materials?

There are a number of researchers looking into creating a suite of systems for the development of computer embedded products which are sympathetic to the designer's mindset and methods, such as Phidgets (Greenberg & Fitchett, 2001; Phidgets Inc.) Voodoo Dolls (Pierce et al., 1999), DTools (Hartman et al., 2005), Switcharoo (Avrahami & Hudson, 2002), Pin and Play (Villars et al., 2005) and Denim (Landay & Myers, 1995). However, these have tended to focus more on the electronics or programming base, whereas we are interested in such systems from a product design angle.

To that effect some of the authors have been involved in the development of low-tech keyboard emulation boxes called IE Units wedded to software building blocks (Gill et. al., 2005b). The IE Units allow rapid prototyping without the usual electronics or programming skills prerequisites and they have been used to empirically measure the performance of real products against physical and virtual prototypes.

The results show that the link between the physical act of holding a product and interaction was more marked than has previously been understood (Evans & Gill, 2006), thus highlighting the need for understanding the precise nature of physicality in the design process. This led to our recent work on physigrams – a diagrammatic notation based on a formal framework for mapping the relationship between physical devices and their corresponding physical actions (Dix et al., 2008) for designers' use.

Case Study

At the Physicality 2007 International Workshop, a design exercise was run in order to explore the influence of different design materials on early design and assess how their physical properties impact issues such as the number and novelty of design ideas and the kinds of designs produced.

Method

The approach taken was open and exploratory rather than controlled, reflecting the aim to uncover new behaviours rather than quantify known ones. A form of ethnographic observation was used that included both live observations and field notes and also video-recording to capture the design process and its outcomes for later review.

Participants were divided into teams of two or three people, and each team was given one kit of design materials to use, namely:

- paper and pencils,
- card and glue, or
- modelling clay, commonly known as plasticine.

Participants were only supposed to use their own materials. Beyond this they were not told how to use the materials, but in fact the materials implicitly suggested ways of use – for example, no team in the paper and pencil group chose to fold or mould the paper to make a model.

In normal design any or all of these materials would be used according to the preferences of the designer at a particular moment. However, in this exercise, teams of participants were given just one kind of material to work with. Thus we were performing something similar to a 'breaching experiment' (Garfinkel, 1967) which deliberately disrupts human activities in order to bring to light aspects that are tacit or taken for granted; although, in standard breaching experiment the conventions broken are social whereas here we are disrupting the ability to choose appropriate materials.

Materials

The materials were chosen to reflect traditional design practice and to cover a range of properties such as, two-dimensional vs. three-dimensional, manual vs. cerebral, malleable vs. constrained.

Pencil and paper are of course used extensively throughout the design process to sketch and work through ideas. Card and glue is an extension of this – allowing very quick rough '3D sketching' to give ideas some shape for discussion, though card can also be used for more refined models. Although "Blue foam" is the most common material used by product designers to create fast 3D models it was not used in this exercise because of the skills, tools and training required as well as accommodation issues (blue foam is very messy and produces fumes when cut with hot wires). Instead, modelling clay was chosen as it has a long history of use in the 3D design process by many designers 'according to the type of product and company practice' (Bordegoni & Cugini, 2005; Verlinden et al., 2001), most notably by the automotive industry (Rekimoto, 1996). It was felt that modelling clay provided the tactile 3D element that blue foam supplied without the need to train the participants.

Design brief and Setting

The brief was to design a hand-held device for producing light that can be turned off and on (see Figure 1). It deliberately kept the technical considerations to a minimum to encourage participants to reflect on the device in relation to the human body.

This exercise looks at the physical design process and its effect on the physicality of the design concept.

We have split you into 3 groups with the same design brief but with different design materials – one will have pencils and paper; the other paper, card, scissors and glue, and the third, plasticine.

We will be recording the process and outcomes of this exercise for later review.

At the end of the exercise we will invite each team to talk about their concept, the process they undertook and the pros/cons of the tools they used.

The brief is to design hand-held device for producing light – it should be able to be turned off & on.

Figure 1. Design brief for participants

The participants were divided in three main groups (one for each material), and each group had 2-3 teams consisting of two or three people of mixed gender. Figure 2 shows the make-up of each group. Each group was given the same design brief but different materials to work with. The teams are labelled team A–H and the participants' names have been anonymised in the transcript fragments presented here[1].

The participants came from various disciplines ranging from computing, arts, design, sociology, philosophy, to human geography and architecture. We had eight teams in all; they were given forty minutes to work on the exercise and then invited to present their concepts and comment on the materials they had used. All the teams were based in the same large meeting room but each team worked independently.

Group	Team	Team members
Card and Glue	A	E^{+}, C^{*}, L^{+}
	B	F^{*}, L^{+}, R^{*}
Plasticine	C	G^{*}, R^{*}, B^{+}
	D	H^{+}, A^{+}
	E	J^{*}, B^{+}, C^{*}
Paper and Pencils	F	A^{*}, B^{+}
	G	D^{*}, E^{+}
	H	H^{*}, F^{*}, G^{+}

<u>Key</u>

* male

$^{+}$ female

Figure 2. Allocation of teams to groups

Initial Observations

The teams varied significantly in terms of their level of exploration. Some teams focused on a single design idea and produced a single prototype, whilst others explored various design ideas and produced a number of prototypes. However, there was no simple relationship

[1] Note the anonymised names are not unique across groups, if there is any ambiguity as to the current group we will refer to a participant as, for example, B>F meaning participant 'F' in team B.

between the groups that used particular materials and them being more prolific or more focused in terms of process and output. Indeed in each group there was at least one focused team and one more exploratory team.

Figure 3. Sample of prototypes

Participants came up with a variety of designs, from fairly traditional functional torches, to a child's bedtime cuddly toy that glowed when stroked. Figure 3 shows some of the prototypes that the participants produced during the design session. We expected that richer materials would lead to more varied designs, however the end picture was far more complex. One team ended up spending most of their time in discussion rather than using the materials as a means to explore designs, and defied the instructions that they should not write things down. It was only at the very end that they used the plasticine to implement an already complete design idea. Another team that only had paper and pencil to work with was most prolific in terms of the number of design ideas they produced. In other cases, the nature of the materials drove the design, so one of the teams working with card ended up producing a card shape design of a torch.

However, while there was not a simple message such as "physicality helps creativity", detailed analysis of the video recordings reveals more subtle patterns of behaviour: including the way groups move between abstract and concrete discussions, the way different groups either comply with or resist the materials they are given, and the complex interactions between the physicality of materials and group dynamics.

Interactions within a team

We will now walk through a single team's interactions step by step. This will give some idea of the way these teams behaved and also begin to highlight issues, which we will discuss more thematically in the final part of the paper.

The team under consideration here is team A, which was supplied with a range of card sizes (A4, A3 and A0 rolls), a glue stick, masking tape and a pair of scissors alongside the design brief. It is interesting to note that there were some remarkable differences between the two teams in the card and glue group (see Figure 2), both in terms of the way the members collaborated and in the type of prototypes they produced. Also, the participants from both teams employed a lot of gestures during the discussion to demonstrate the ideas they were trying to get across.

The participants in team A were ‘E’, ‘C’ and ‘L’. The team members initially spent a substantial time discussing the design concept and exploring various alternatives with the materials they were given. They started with the obvious idea of a torch as a hand-held device that can be switched on and off, but they soon moved away from that concept, as ‘C’ later confirmed in the presentation session.

> A>C: we started with the obvious torch, you just press a button to light... we thought that was very boring …

They narrowed their design focus by thinking of a possible scenario that they may need such a device for.

> A>C: … we thought about what you need light for, we came up with the very plausible scenario of you wanting to read under your blanket without disturbing or being disturbed

They went on to explore the shape of the device. ‘E’ starts by naturally rolling the card into a cylindrical shape, but ‘L’ suggests, “what if it can be a handheld device, that is really flat and you can unfold and keep in your pocket”.

The type of light source was the next issue that was discussed.

> A>C: something that illuminates like a keyhole
>
> A>L: It can be any light, a strobe light ... an instant strobe light
>
> A>E: a head torch?
>
> A>L: what about something rechargeable, that’s not very heavy

The team members carried on with their discussion until their ideas started getting clearer. They then proceeded to make some prototypes from the materials they were given and ended up producing three main prototypes, one based on each team member’s design concept. However, there was some degree of collaboration between the team members during the model-making process, as described below. Their aim was to produce a reading light that is inconspicuous and more importantly, does not look like a traditional reading light.

Prototype 1

‘C’ starts by rolling up an A4 size card and taping it with some masking tape to make a tube. ‘E’ is quick to point out the issue with using a straight tube in a tent.

> A>E: In a tent, the problem is that with a straight tube light (demonstrates using the glue stick) the angle is wrong, what you want is a torch that bends as an upside down u shape on the top (demonstrates using gestures)

'E' proceeds to make a prototype (see Figure 4) by using an A4 size card (4a), she cuts angles off with the pair of scissors (4b) and uses the masking tape to join the edges (4c). She cuts open the end to show where the light shines out from (4d). She later improves the design by adding another piece onto the end (4e).

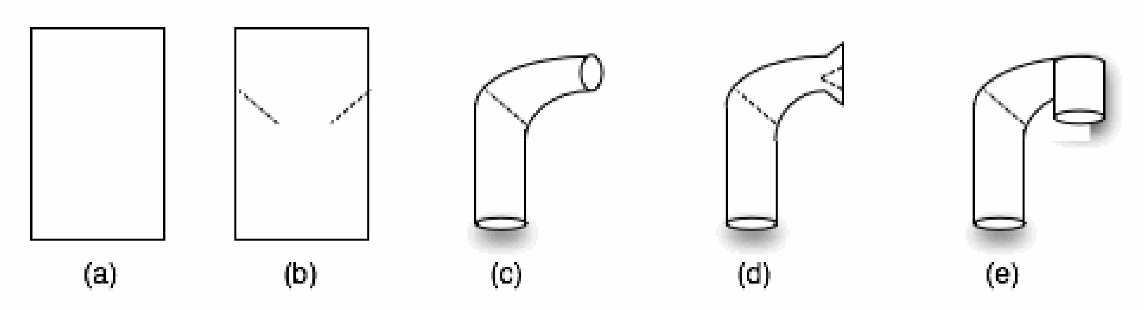


Figure 4. Various stages of prototype 1

During the presentation, 'E' describes her thoughts behind her prototype as follows:

> A>E: a lamp, the same idea as reading in a tent or reading under the blanket, you are able to hold it in the hand, you want the light to only go on the book, but not on the cover, so you don't get caught by your parents… and then it would be good if it's heavy at the bottom so it doesn't, so you don't need to hold it

Prototype 2

'C' rolls out an A4 size card into a thin tube to produce a reading light that can be attached to the outside of book, which he demonstrates by folding a piece of card to represent the book (see Figure 5). 'L' suggests having a flexible light at the top so one can easily point to different places on the page but 'C' remarks that, "a 'V' shaped is better for shedding even light across the whole page". 'C' adds, "it is better to stick it outside the card otherwise you can't flip the pages of the book".

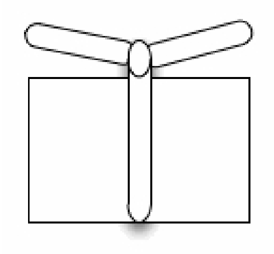


Figure 5. Example prototype 2

They then engage in a discussion as to where the batteries should be fitted. 'C' suggests, "on the flat side (showing the book spine)", hence the need for "flat batteries".

'C' later describes his prototype as follows:

> A>C: a book reading lamp for underneath your blanket, you basically hold it behind your book, and there is a little switch here (imaginary one)… you move it up and down, it shines light on your page, you can flip the pages without the light being in your way, and you can hold it in one hand and hold up the blanket with the other hand

Prototype 3

'L' scrunches a piece of card to demonstrate an 'ergonomic handheld' shape that can be shone onto things. She engages her team members by suggesting:

> A>L: imagine if it is made out of rubbery material (holding the scrunched up paper) that you can squeeze and the best thing is that it doesn't look like a torch, so if you get caught you don't have a torch, you have a stress ball or something!

'E' suggests, "a teddy bear on the wrist", 'L' embraces that idea.

> A>L: oh yes a teddy bear, so the light comes out of its eyes ... how do you make that ... ok I try to build that one now.

'L' starts by making a pattern for the teddy bear. 'E' and 'C' join in with their own teddy bear models made from different size cards. 'E' cuts out a fairly large flat teddy bear. 'C' instead uses the ergonomic shape that 'L' made earlier as a mould and adds layers to it, resulting in an amorphous hand-held shape.

'L' however works towards making a medium size 3D teddy bear model. She makes the 3D body shape from two layers of cards cut-outs and stick the cut-out angles together to produce a 3D effect. She adds some scrunched up card as stuffing between the layers. She then proceeds to make a 3D head shape with some stuffing in between. Before joining the head to the body, L adds some flat arms and secures everything with the masking tape.

Figure 6. Prototype 3

Looking at the end result (see Figure 6), 'L' remarks, "it is quite a big teddy lamp!"; 'C' calls it "iTeddy".

'E' adds a belly button to her flat teddy model and suggests, "this can be pushed, like that (holding the teddy up)".

'L' sticks a button on her 3D teddy model too but 'C' remarks 'I don't think we need the button, in a way, you can just squeeze it (demonstrates using the model) to switch it on". 'L' agrees and removes the button and says, "keep it as a conceptual sketch ... and how its gonna look, the light is gonna come out (squeezing the teddy)"

'L' later explains the thoughts behind her design:

> A>L: when I was little I used to take my microscope lamp to read, it wasn't easy to switch off. So I thought of something that was hand shaped that you could squeeze... and then we did this one (showing the teddy bear) and then thought it would be cool, if your parents actually catch you and you don't have a lamp in your hands, so you

have like a teddy bear... you just press the teddy and the light shines …we thought a book and it would be a good supplement for a children's magazine.

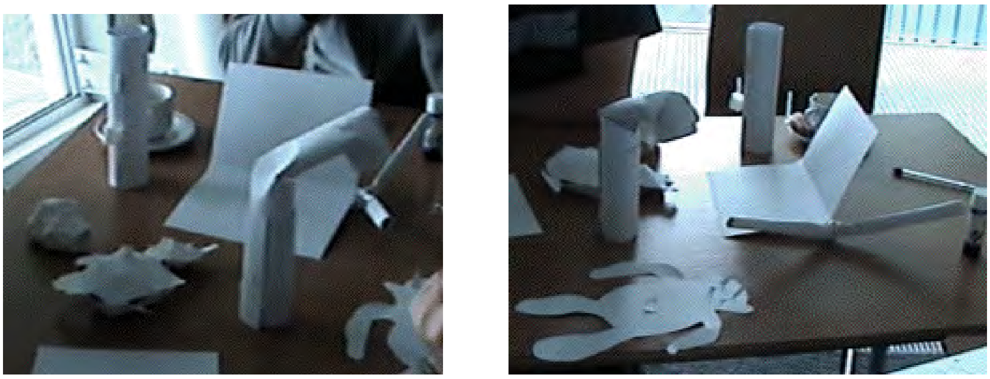


Figure 7. Some prototypes made by Team A

Figure 7 shows a selection of the prototypes that Team A produced.

Thematic Analysis

The initial ethnographic observations and field notes carried out on site were later supplemented by a detailed transcription of the video recordings. Our data highlighted several issues and through an in-depth iterative analysis, we identified individual topics and activities as well as a number of recurrent themes.

Even within this one team, we can see a wide variety of behaviours: from designs driven by the physical properties of the materials, for instance when 'C' rolls up a piece card in prototype 1 and 'finds' a classic cylindrical torch shape, to more abstract discussions of properties:

A>L: what about something rechargeable …

There were also some underlying trends, for example, the groups with paper and pencil tended to produce more fragments of ideas, but not necessarily more finished design concepts. However, the picture is typically more complex, and the themes have helped us to unpick the rich interplay between materials, design brief, team makeup and dynamics.

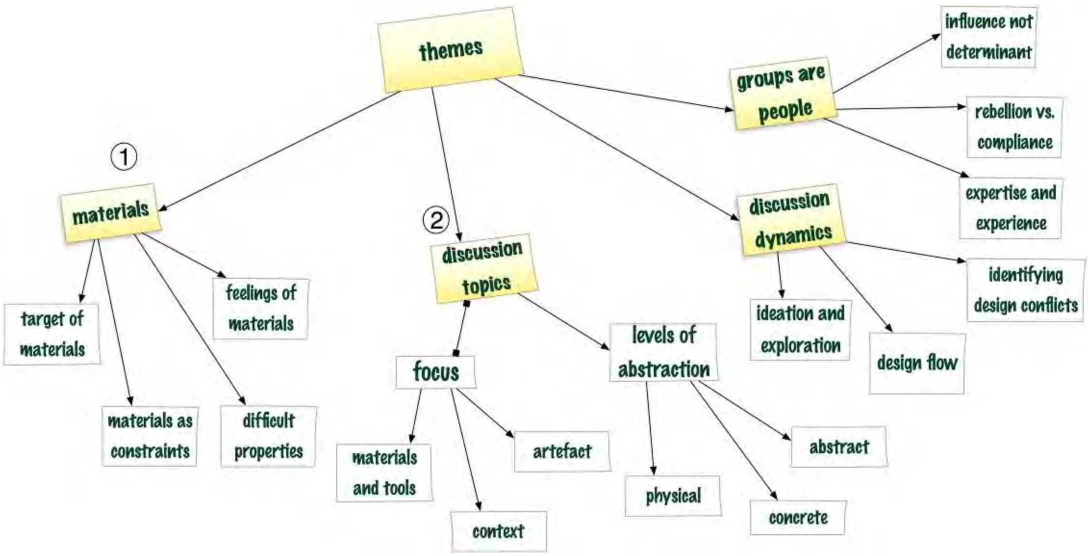


Figure 8. Breakdown of themes

We have categorised the themes into four main classes (see Figure 8):

① those relating specifically to what the materials give to the teams;

② those concerned with the topic of the design discussions;

③ those relating to the flow or dynamics of that design discussion; and finally

④ those relating to the personal and interpersonal factors within the group.

We will briefly discuss the first two of these themes (① and ②) as they explicitly bring out issues related to the physical properties of the materials during the early design stage. We also support this discussion with excerpts from the transcripts.

Materials

Target of materials: prototype – design – product

Some of the materials given to the groups were clearly intended to be used as a part of the prototype itself: card, plasticine, whilst others, such as the paper and pencil, are used during the design process, but are not evident in the prototype itself (except insofar as the design sketches are on paper.)

It was interesting to note that none of the teams given only paper and pencil used this to create a paper model by folding, tearing etc., despite having ample paper to draw on and to play with. It seems that once participants regarded the paper as a thing for drawing on (design material), it became impossible to see it as a raw material for construction. This conservatism was seen despite many other forms of challenging or subversive behaviour, thus suggesting that it is a very hard mindset to change.

Occasionally, the discussion turned to the actual materials that would be used on the product assuming the design were realised, for example,

> C>G: "… and LumaTed is made from luminous material -"Philips Lumalight", with a sort of translucent material, that can light up in different colours, and you interact with it by stroking, so it will have a set of capacitor sensors that will allow you to stoke it in different ways and the more you stroke it the brighter it gets...

This group even went on to give LumaTed a price. However, few groups explicitly differentiated the prototype material from the production material, probably exacerbating some team's tendency to be 'trapped' by the materials (see discussion below).

The design materials included things like the scissors for cutting, but these were sometimes used in unexpected ways. For example, team B considered using a roll of masking tape to draw smaller circles, and the scissors as a compass to draw wider ones using the point as a pivot. Furthermore, materials were often drawn in from the environment. Team D used a water bottle extensively both to inspire their design and eventually as part of their prototype (that is both as a design material and a prototype material). Another team rolled plasticine on the rough walls in an attempt to produce textured surfaces.

Ethnographies of group activities in other domains have demonstrated the importance of shared artefacts in coordinating actions; for example, the large common display in Heath and Luff's (1992) analysis of the London Underground. In our study we too found that the physical nature of the design materials is used to manage sharing and provide shared focus. For example, in Team F, members 'A' and 'B' started sketching ideas on separate sheets of paper, but rapidly switched to using the same sheet of paper, thus using the paper as a shared artefact to maintain collaboration and generate design ideas.

Materials as constraints

Constraints can sometimes be seen as a bad thing: limiting, holding back. However, psychological research on creativity and problem solving has often found that constraints can inspire creative designs (Ormerod, 2002), partly because they focus the design and partly because they sometimes prevent the 'obvious' solutions.

As we saw earlier during the production of Prototype 3 (Figure 6), L was did not let the card, an essentially two-dimensional material, hinder her design for a 3D teddy bear model. In fact her ability to produce such a refined model most likely reflects her background in fashion design.

However, most of the other participants often referred to the fact that they would have preferred one of the other materials, sometimes during specific parts of the design process. Some wanted pens to sketch with or other set of materials to use, for instance with Team A:

> A>C: yes, I mean, I think we were lacking some pen to scribble

> A>L: or plasticine!

and Team B:

> B>L: I think to sketch out an initial design probably would've been handy, or a way to, sort of, come up with some ideas

Similarly, Team H were frustrated that they have not been given clay to try out their ideas:

> H>F: I think… the sketching, or the pencil and paper were ok for the initial communication of ideas and... summarising what we thought was right…

> H>G: Yeh, it would've been good to have something that we could actually mould and actually get more feeling about the actual prototype we came up

Noticeably there was a greater tendency to ask for clay than card. This may reflect its greater malleability, for example Team A admitted during their presentation that they found the card “difficult to bend to the shapes that our minds had formed in our head” and Team F "… tried to look at some sort of more organic bioforms, shapes, but paper is not a very good medium for doing that…". The popularity of the clay may well also be because it was regarded as being more 'fun' (see later).

Teams also responded more subtly to their materials: the majority of card-based prototypes were formed from cylinders and other rollable shapes. However, this material did not totally determine the design, as we saw Prototype 3 (Figure 6) included scrunching up the paper, making it in effect more like the clay.

Difficult properties

Some properties that were mentioned during design discussions were difficult to recreate in any physical prototypes. This included the weight of objects (hard with card), softness (hard with plasticine), the light itself.

> F>A: how do we produce light?... is a spark a light? …

It is interesting to note that only the teams using paper and pencil discussed energy and 'light' at length. This is perhaps related to the fact that they engaged in more abstract discussions and that whilst weight, energy and light are very physical in one sense, they are also somewhat ethereal properties. But even this was not a universal rule as Team E used a cone of paper to " simulate light".

Feelings of materials

Whilst the materials did not determine team behaviour, the fact that plasticine was a child's toy certainly seemed to influence the teams' attitudes. The teams using plasticine appeared to operate in a more playful and sometimes wacky manner; for example, at one point team E produced a Petri dish. Team C spent a period discussing ideas, but all the while each holding a piece of plasticine, kneading and playing with it, but not using it to make anything. Even during the discussion stage, C>G starts playing with the plasticine and says "Oh yeh, well we really enjoyed the plasticine! And yeh that was fun".

In contrast, card suggested more formal/serious designs:

> B>R: In terms of our process it was very much orientated to what we thought we could do with the materials we'd been given …we feel that it was good for making something that was solid, if you drop it, it probably won't break, but more than that it's not very expressive …

Discussion topics

The topics that were discussed by the teams were based on different levels of abstractions (from physical, concrete to abstract) and focus (on the artefact, the context and the materials). Some teams spent more time in one or other kind of discussion, but also each team moved between kinds, at one moment discussing concrete design ideas, at another, more abstract discussion of requirements.

Level of abstraction: physical – concrete – abstract

Some of the discussion focused around the physical nature of the materials and models that were in their hands. For example, team B focused on making a torch that was as realistic as possible:

> B>R: … most of our process was about making a model that looks relatively realistic, or at least as realistic as we could get.

They also used physical things in the environment (such as the water bottle mentioned previously) to augment their design or to demonstrate or stimulate ideas:

> G>D: So the idea was a watch with light... so the concept was this watch (showing his watch on his wrist), putting some lamps inside the clock/the watch, by sensors, touching it you can make it work

and even their own bodies:

> H>F: yeh sure... you are just cupping the light... (demonstrates two different ways of cupping using his hands)

> H>F: ...I think it would be great... for warming light... (rubbing hands together)…

Note, team H was in the paper and pencil group, so had no obvious prototyping material to create this sort of physical focus for their discussion.

At other times, the discussion was still quite specific, talking about a particular design or scenario of use, but without having it physically to hand. For example, team C were refining the shape of their teddy bear:

> C>G: Yeh to show... we have a version... when the child grabs, cuddles it, it will come on, that's one situation, or strokes it... we have yours where it's at the end of the bed and we have to look for it, so we have to ask it, call its name and it'll come on …

Finally, there were times when the discussion was at a more abstract level discussing general ideas, properties or dimensions. For example team H discussed ideas of "discrete feedback" and needing some form of "discrete interaction" and Team B considered the 'primitives' afforded by their material (card):

> B>R: we've got a circle (holding the masking tape) so we can use that to create a precise circle, we can mark it with that but we can score the card, so those are the primitives…

There was a tendency for the paper and pencil group to have more abstract discussions, but this is far from being their preserve and many teams engaged in some form of more abstract discussion. What was evident was that at the point at which discussions became more abstract the teams with prototyping materials 'stepped back' from the materials … and one plasticine team even 'cheated' and used paper and pen!

Despite this 'stepping back', this is not to suggest that these different levels of abstraction are independent discussions. On the contrary, there is a constant interplay where more abstract discussions lead back to concrete design suggestions:

> F>A: … produces light and there are implications with that… it has a battery, it has a bulb – that's the normal way to produce light, although they could have one of those, err, wind up ones...

or lead to physical design solutions:

> H> G: I guess if you're looking for discrete… what you need is some sort of discrete interaction like clapping …

and even physical on concrete considerations prompt generalisations.

> H>F: (sketching on paper) I think if you look at the fire, there's a couple of things that, err, you can read from the physicality of the fire.. if you place more logs on the fire, you can see how long the fire might eventually burn... and if the fire dims, you see the flames going down, you see that you have to put more on the fire...

Focus: artefact – context – materials and tools

From the quotes and examples we have seen so far it is evident that the discussion topics sometimes focused on the artefacts that are being designed:

> F>C: we started with the obvious torch

sometimes on the context in which the artefact would be used:

> A>C: … we thought about what you need light for, we came up with the very plausible scenario of you wanting to read under your blanket without disturbing or being disturbed

and sometimes on the materials, tools or process of design itself:

> E>B: …we used all kinds of tools, we split a pen apart and used the top and cut with this and used round things as forms and used the paper to simulate light and we borrowed some of the green plasticine from the other group...

Each of these could be discussed at each level of abstraction as exemplified in Figure 9. We have seen several instances of these, such as in team A's prototype 2, they not only made the reading light, but they also used a folded piece of card to simulate a book: a physical model of the context.

Some parts of the picture are more common than others; for example, most of those teams with physical materials spent a considerable amount of time manipulating the physical artefacts. However, as previously noted, the abstract parts of this space are not the preserve of the paper and pencil group only, indeed team C, in the plasticine group, at one point raised the following:

> C>G:… And we decided to focus it on children. And it seemed reasonable in that case to try to make it into some kind of night light, something that will help children when they're feeling frightened at night. And so from that we got a set of properties that we thought we would want to express through this, we wanted to help children feel safe and secure at night, it would be something that would be easy to interact with when they're kind of in that semi-wakeful state, something that would be soft, or warm, smooth, stable, robust... and out of all that came "LumaTed" ...

The above excerpt illustrates the flow of the discussion which starts off in abstract context "focus on children", moves to concrete artefact "night light", then back to abstract context "feeling frightened at night", to abstract artefact "set of properties", and eventually back to more refined concrete artefact "LumaTed".

Focus / Level of Abstraction	artefact	context	materials/ tools
physical	holding torch	making book	rolling the card
concrete	mentions torch	scenario	scissors as compass
abstract	need to be rigid	requirements	'primitives'

Figure 9. Level of Abstraction vs. Focus of discussion topics

Conclusion

Our case study showed that a minimal design brief with fairly low-tech materials can in fact generate a wealth of information, thus reaffirming the importance of producing low fidelity prototypes at an early stage in the design. Although we set out to explore the role that physicality plays in the design process, the results defy simplistic conclusions such as "physicality promotes creativity", or even the opposite.

Our in-depth thematic analysis however reveals dimensions along which general trends can be seen. For example, the tendency for the teams with paper and pencil (typically) to engage in more abstract discussion, is probably one of the reasons for the greater number of (often fragmentary) design ideas. But again this is not as simple a story as stating that teams with prototyping materials tended to do such and such things. Although we have focussed on the materials and discussion topics, we cannot ignore the effects that the flow or dynamics of the discussion, and the personal and interpersonal factors within a group have on the design process.

The way that materials were utilised was partially a consequence of preconceptions brought into play by the backgrounds of the participants. So even within a single group, card is treated as if it were clay (crumpled to conform with the shape of a clenched hand), as an essentially two dimensional material (forming a bear as a cut out) and finally as a textile (forming a three dimensional teddy). So while the materials supplied may influence the output, it was also clear that the experience the user brought to the table partially influenced the design.

Individuals within groups and the way they worked together, would often mean they defied the restrictions or paths suggested by their materials – including rebelling completely, as with the plasticine team who, against the rules, got paper and pencil. This ability to move against the natural tendencies of physical materials seems very dependent on the characters of individuals and teams.

For practicing designers, but even more so for students, this does prompt questions as to how to maximise the benefits of specific physical materials in prompting new ideas, whilst also at appropriate moments during design activity, to 'escape' the practical and cognitive limitations they create. We do not answer this question here, but believe that the rich understanding of the design process we have produced is a step towards this.

Much of the more theoretical understanding of physical artefacts is focused on objects that concretely achieve physical goals: for Heidegger the way a hammer is 'ready to hand' in the act of joining wood with nails, or for Gibson, the way a rock if of suitable size 'affords' sitting upon. In the case of objects in a 'natural' (pre-technological) environment, Gibson argues that if we are well adapted to the environment, then our perceptions are tuned so that the affordances are immediately perceived; we are creatures tuned for action. However, as soon as we consider technological objects, things become more complex. Even turning a door handle needs to be considered as a sequentially unfolding chain of learnt associations and skills, as well as more immediate visual and haptic perceptions (Gaver, 1991). Similarly, 'ready to hand', while frequently misquoted, is not a matter of 'walk up and use' but is the product of culture and skill.

In this paper, we have looked at materials in design - that is physical objects that are for essentially cognitive tasks. What a material 'affords' under such circumstances is even more finely dependent on the past knowledge and skills of those using them (e.g. fashion designer vs. sculptor); and yet the material is not entirely open, without influence, like a piece of wood being carved, it has a grain, a set of uses that are easier than others, that fall more readily to hand or mind. Building an adequate practical and theoretical understanding of such a nuanced and context sensitive area is no easy task, and one we have by no means accomplished, but is, we believe, a valuable goal.

Acknowledgements

We would like to thank all participants of the Physicality 2007 workshop for their contribution and invaluable insights they provided us by taking part in the design session. This work has been supported by the AHRC/EPSRC funded project “DEPtH: Designing for Physicality” (http://www.physicality.org/) as part of the “Designing for the 21st Century” research initiative (http://www.design21.dundee.ac.uk/).

References

Avrahami D and Hudson SE. (2002) Forming Interactivity: A Tool for Rapid Prototyping of Physical Interactive Products. In Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS’02 (pp. 141–146). New York, ACM Press. http://doi.acm.org/10.1145/778712.778735

Baxter, M. (2002). Product Design: A Practical Guide to Systematic Methods of New Product Development (pp. 27-28). Cheltenham: Nelson Thornes (Publishers) Ltd.

Bordegoni, M. and Cugini, U. (2005) Design products with your hands. In Proceedings of Virtual Concept 2005. Biarritz, 8-10 November 2005.

Branham, R. (2000). Given the radically changing work environment and new worldviews, what kinds of new 'tools' do designers need to survive and successfully deal with tomorrow's design problems? In Proceedings of IDSA 2000 National Education Conference. Louisiana, USA.

Buchenau, M. and Suri, J.F. (2000) Experience Prototyping. In Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques (pp. 424 – 433). New York, USA.

Clark, A. (1998) Being There: Putting Brain, Body, and World Together Again. MIT Press.

Dix, A., Ghazali, M., Gill, S., Hare, J. and Ramduny-Ellis, D. (2008) Physigrams: Modelling Devices for Natural Interaction, under consideration for publication in Formal Aspects of Computing.

Evans, M. and Gill, S., (2006) Rapid Development of Information Appliances. In Proceedings of International Design Conference Design 2006. Croatia, 15–18 May 2006.

Norman, D. A. (1998). The Invisible Computer: Why good products can fail, the personal computer is so complex and Information Appliances are the solution (pp189). London; Cambridge, MA, MIT Press.

Garfinkel, H. (1967) Studies in Ethnomethodology. Englewood Cliffs, NJ: Prentice- Hall.

Gaver, W. (1991) Technology affordances. In Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology (pp. 79 - 84). New Orleans, Louisiana, US, ACM Press.

Ghazali, M. and Dix, A. (2005) Knowledge of Today for the Design of Tomorrow. In Proceedings of the 2nd International Design and Engagibility Conference (IDEC 2005). Edinburgh, 6th Sept. 2005.

Gibson, J.J. (1979). The Theory of Affordances. In R. Shaw and J. Bransford (Eds,) Perceving, acting and knowing: Towards an ecological psychology (pp. 67-82). Hillsdale, NJ: Erlbaum.

Gill, S., Johnson, P., Dale, J., Loudon, G., Hewett, B. and Barham, G. (2005) The Traditional Design Process Versus a New Design Methodology: a Comparative Case Study of a Rapidly Designed Information Appliance. In Proceedings of the Human Oriented Informatics & Telematics Conference. University Of York, UK.

Gill, S., Loudon, G., Hewett, B. and Barham, G. (2005b) How to Design and Prototype an Information Appliance in 24 Hours – Integrating Product & Interface Design Processes. In Proceedings of the 6th International conference on Computer Aided Industrial Design and Concept Design conference, University of Delft, The Netherlands.

Greenberg, S. and Fitchett, C. (2001) Phidgets: easy development of physical interfaces through physical widgets. In Proceedings of the 14th Annual ACM Symposium on User interface Software and Technology, UIST '01 (pp. 209-218). Orlando, Florida, ACM Press. http://doi.acm.org/10.1145/502348.502388

Hartman, B., S. Klemmer, M. Bernstein and N. Mehta (2005) d.tools: Visually Prototyping Physical UIs through State-charts, in Extended Abstracts of UIST 2005, ACM Press.

Heath, C. & Luff, P. (1992). Collaboration and control: Crisis management and multimedia technology in London Underground line control rooms. Computer Supported Cooperative Work, 1 69–94.

Hutchins, E. (1995) Cognition in the Wild. MIT Press, Cambridge, MA.

Landay, J. and B. Myers. (1995) Interactive sketching for the early stages of user interface design. In I. R. Katz, R. Mack, L. Marks, M. B. Rosson, and J. Nielsen (Eds.), Proceedings of the

SIGCHI Conference on Human Factors in Computing Systems (pp. 43–50). Denver, Colorado, ACM Press/Addison-Wesley Publishing Co. http://doi.acm.org/10.1145/223904.223910

Ormerod, T., MacGregor, J., and Chronicle, E. (2002). Dynamics and constraints in insight problem solving. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(4), 791-799.

Pierce, J.S., Steams, B.C., and Pausch, R. (1999) Voodoo Dolls: Seamless Interaction at Multiple Scales in Virtual Environments. In Proceedings of the 1999 Symposium on Interactive 3D Graphics (pp.141-145).

Phidgets Inc. http://www.phidgets.com/

Ramduny-Ellis, D., Dix, A., Hare, J. and Gill, S. (2007) Proceedings of the Second International Workshop on Physicality, Physicality 2007 (2-3 Sep. 2007). Lancaster University, UK, ISBN 978-1-905617-60-9, UWIC Press. http://www.physicality.org/Physicality2007

Rekimoto, J. (1996) Transvision: A hand-held augmented reality system for collaborative design. In International Conference on Virtual Systems and Multimedia (VSMM'96) (pp. 85-90).

Suchman, L. (1987) Plans and Situated Actions: The problem of human–machine communication. Cambridge University Press.

Snyder, C. (2003), Paper Prototyping, Morgan Kaufmann, San Francisco, CA.

Verlinden, J.C., T. Wiegers, H. Vogelaar, I. Horvath, J.S.M. Vergeest (2001) Exploring conceptual design using a clay-based wizard of Oz technique. In: G. Jahannsen (Eds.), 8th IFAC/IFIP/FORS/IEA symposium on analysis, design and evaluation of human-machine systems (pp. 211-216), VDI 2001, Kassel, September 18-20, 2001.

Villar, N., Lindsay, A. and Gellersen, H. (2005) Pin & Play & Perform; A rearrangeable interface for musical composition and performance. In Proceedings of the 2005 conference on New interfaces for musical expression. Vancouver, Canada.

Wheeler, M. (2005) Reconstructing the Cognitive World: The Next Step. Bradford Books.

Devina Ramduny-Ellis

A researcher in the Computing Department at Lancaster University in UK. She started her research in the area of Human Computer Interaction and Computer Supported Collaborative Work and was awarded her PhD in May 2003 for her work on developing architectural frameworks for enhancing temporal interface behaviour for distributed collaborative users. Her research interests ranges from devising methods for work analysis and interaction properties, modelling techniques and interface design, to computing infrastructure and employing empirical field studies involving ethnographic and other qualitative techniques to inform novel designs and prototypes. She has worked in several research projects within an academic environment, which covered a wide range of areas including ubiquitous computing, dependability, decision management and interfaces and infrastructure. She has over 25 publications and she is currently managing and working on the DEPtH: Designing for Physicality AHRC/EPSRC funded project. She has been actively involved in starting up a multi-disciplinary community in physicality through the Physicality 2006 and Physicality 2007 International Workshops.

Joanna Hare

Jo graduated in 2004 from De Montfort University, Leicester with a BSc. honours degree in Industrial Design, during which she completed a placement year working within BMW's Design and Development building in Munich. After graduating, Jo worked for Exmoor Plastics, a medical device manufacturing company in Taunton, in conjunction with PDR as part of the Knowledge Transfer Partnership (KTP) scheme. The KTP project focused on increasing the company's productivity by decreasing product development time. New procedures were successfully implemented that defined and selected new project proposals to take forward. Jo is currently working within the PAIPR research group on Design for Physicality (DEPtH), an AHRC/EPSRC funded project under the Design for the 21st Century programme.

Alan Dix

A Professor in the Department of Computing at Lancaster University in the UK. He was a mathematician by training, and mathematics is still his first love, but he has worked in Human-Computer Interaction since 1984, has published over 300 articles and is author of a key textbook in the area. He has worked in several universities, agricultural engineering research, local government and hi-tech start-ups. His interests are eclectic: formalisation and design, physicality and digitality, the economics of information, structure and creativity and the modelling of dreams. Recently he and a colleague have developed technology for autonomous pixels that can be configured in turn any surface or space into a two or three dimensional display.

Steve Gill

A product designer with 15 years experience in industry and academia. He is Director of Research of the Cardiff School of Art & Design and Director of the Programme for Advanced Interactive Prototype Research (PAIPR) within The National Centre for Product Design & Development Research (PDR). He has designed or product managed around 50 products to market and has published 20 academic journal and conference papers. Steve has a range of research interests related to product design and development including the rapid design development of information appliances, the role of physicality in the design process and the implications for product design development of testing early information appliance prototypes within their context of use. He works closely with academic partners such as the University of Lancaster and those in blue chip industry such as UIQ and Sony-Ericsson (whose Smartphone division have altered their interface design strategy to implement his research work through a working partnership).

The Design Journal VOLUME 13, ISSUE 1 PP 48–76 REPRINTS AVAILABLE DIRECTLY FROM THE PUBLISHERS PHOTOCOPYING PERMITTED BY LICENSE ONLY

Physicality in Design: An Exploration

Devina Ramduny-Ellis, Alan Dix and Martyn Evans
Lancaster University, UK

Jo Hare and Steve Gill
Cardiff School of Art & Design & The National Centre for Product Design & Development Research (PDR), University of Wales Institute, Cardiff (UWIC), UK

ABSTRACT Both the nature of many products and their process of creation are becoming increasingly digitally mediated. However, our bodies and minds are naturally conceived to interact with the physical, so crucial design information can be elicited by constructing meaningful prototypes. This paper examines how physical materials impact early design through a study that explores how groups with very different materials tackle a common design challenge. The inherent physical properties of the materials and the ways in which designers interpret and manipulate them give rise to subtle patterns of behaviour.

DOI: 10.2752/146069210X12580336766365

These include the ways in which groups move between abstract and concrete discussions, the way groups comply with or resist the materials they are given, and the complex interactions between the physicality of materials and the group dynamics. This understanding is contributing to our research in explicating the fundamental role of physicality in the design of hybrid physical and digital artefacts.

KEYWORDS: physicality, design process, prototyping

Introduction

Design has a long tradition of artistic engagement with, and manipulation of, materials as an intrinsic part of the design process. A material-centric approach, exemplified by the artisan designer, has shaped our modern understanding of design (Pevsner, 1991; Raizman, 2004; Woodham, 1997). Potter (1969) describes the activities of the artisan designer as being driven by the manipulation of materials in workshops where they 'get their hands dirty'. Traditionally, product design has focused on artefacts designed using physical materials such as clay, wood, metal or plastic (Heskett, 1980). The human condition is made palpable by the material world it inhabits and meaning is taken from physical objects that colour our everyday lives (Dant, 1999; Woodward, 2007). We have a propensity for the physical materials we have experienced throughout our lives and understand the behaviour of stone and wood, water and metal, for example, as being natural to us. However, this is changing as manmade artefacts increasingly include digital elements, from MP3 players to mobile phones, and in the design process itself, sketches and clay models are giving way to Computer-aided Design (CAD) and virtual representations.

It is in this context that we are seeking to explicate the properties of physical materials and physical artefacts and the way we understand and manipulate them, so that we can firstly, better inform the design of hybrid digital/physical artefacts and secondly, understand the impact of changing tools and techniques on the design process.

The authors' previous work has considered properties of commercially available electronic and domestic products in order to uncover the ways in which designers exploit the physical nature and placement of interface controls. Quite subtle differences have a major influence on the naturalness of interaction for end-users (Ghazali and Dix, 2005), that is the extent to which the product exploits the user's automatic or subconscious reactions and behaviours. Another study of mobile phone prototypes showed that physical mock-ups of the interfaces can generate significantly more useful user feedback compared with purely on-screen interactive prototypes; whilst this was as expected, more surprisingly, even quite low fidelity mock-ups were as useful as high fidelity ones (Gill *et al*, 2005a).

In this paper, our focus is on the role of physicality within the design process. We investigate the manner in which material properties affect both the process and outcomes of design activity. Our study presents the findings of a design exercise based on using a variety of materials to explore a common design challenge, which was undertaken at *Physicality 2007*, the Second International Workshop on Physicality (Ramduny-Ellis *et al*, 2007). A detailed analysis of the design exercise identified several recurrent themes and issues regarding the rich interplay between materials, design brief, team make-up and dynamics. This thematic discussion is addressed in the latter stages of this paper. Understanding the fundamental role of physicality in product design has been a key motivation underpinning our research.

Motivation

With advancements in technology, the form of objects is no longer driven by the technologies within them (Evans and Sommerville, 2007). Traditional modes of understanding for the product's expression of meaning no longer apply (Norman, 1998a; Vihma, 1995). Human-centred development of computer embedded products, and more specifically user targeted 'information appliances' (products embedded with computers such as mobile phones, digital cameras) are at the crossroads of a number of disciplines (Norman, 1998b); therefore their development process can frequently be disjointed with the physical and digital elements being designed in isolation. These physical and digital processes are often only combined for user testing towards the end of the development process when major design changes are impractical. Baxter (2002, pp. 27–28) and Branham (2000) identify the need for new tools to overcome this problem.

In order to create an effective and a pleasurable experience for the user, designers need to ensure that the relationship between the physical and digital form is explored thoroughly at the early stages of the design process when ideas can be worked through quickly (Rakers, 2001). One of the ways that this could be achieved is by building prototypes, a true to life model of the design in progress, equipped with some properties. Ulrich and Eppinger (2003) define prototypes as an approximation of the product on one or more dimensions of interest, which serve as tools for learning, communication, integration and milestones.

Prototyping techniques such as Experience Prototyping (Buchenau and Suri, 2000) and Paper Prototyping (Snyder, 2003) go some way towards exploring the physical and digital form, particularly in their inclusion of physicality. However, methods that retain their 'quick and dirty' hands-on approach while incorporating more accurate simulation are still required. One of the problems facing designers in modern businesses is that the pressure to deliver to very tight deadlines and within tight profit margins results in physical prototypes being rarely produced, or if they are built, often having limited functionality. The ability to build meaningful prototypes without investing large amounts of time and money is key to obtaining significant design information from product and user testing.

Increasingly, the products we use are a synthesis of digital and physical elements and, for the user, these become indistinguishable. As hybrid physical/digital products are developed, designers have to understand what is lost or confused by this added digitality – and so need to understand physicality more clearly than before.

Our concern with the nature of the physical world and our interaction with it is not new; it has been a concern for philosophers for many years, most notably Heidegger (1962), and is the topic of ongoing discourse, particularly related to issues of the embodied mind (Clark, 1998; Wheeler, 2005). For some within psychology, the traditional 'inside-to-outside' Cartesian conceptions of cognition have given way to an increasing acceptance of the importance of physical embodiment for cognition. This is explicit in frameworks such as distributed cognition (Hutchins, 1995), where the role of physical artefacts and multiple actors is seen as essential for 'cognition' to occur and also in the concept of situated action in Suchman's early work at Xerox (Suchman, 1987), which was seminal in bringing ethnographic approaches into interaction research and practice. Environmental psychologists such as Gibson have also explored this area and Gibson's concept of affordance, the ways in which an object is fitted for human action, has entered the vocabulary of interaction design (Gibson, 1979).

There is much research activity looking into creating systems for the development of computer embedded products which are sympathetic to the designer's mindset and methods. These include Phidgets (Greenberg and Fitchett, 2001; Phidgets Inc., 2008), Voodoo Dolls (Pierce *et al*, 1999), DTools (Hartmann *et al*, 2006), Switcharoo (Avrahami and Hudson, 2002), Pin and Play (Villar *et al*, 2005) and Denim (Landay and Myers, 1995). However, these have tended to focus more on the electronics or programming base, whereas we are interested in such systems from a product design perspective.

The authors have developed low-tech keyboard emulation boxes (*IE Units*) which link to software building blocks (Gill *et al*, 2005b). The *IE Units* allow rapid prototyping without the usual electronics or programming prerequisites and have been used to empirically measure the performance of real products against physical and virtual prototypes. The results show that the link between the physical act of holding a product and interaction was more marked than has previously been understood (Gill *et al*, 2008), thus highlighting the need for understanding the precise nature of physicality in the design process. This led to our recent work on physigrams – a diagrammatic notation based on a formal framework for mapping the relationship between physical devices and their corresponding physical actions for designers' use (Dix *et al*, 2009). Having recognized that physicality is important in devices and prototypes, this paper looks at how the physicality of design materials may affect the design itself.

Case Study

At the *Physicality 2007* International Workshop held at Lancaster University, a design exercise was undertaken to explore the influence of a variety of materials in the initial stages of the design activity and assess how their physical properties impact issues such as the number and novelty of design ideas and the kinds of designs produced.

Method

Our approach was open and exploratory rather than controlled, reflecting the aim to understand new behaviours rather than quantify known ones. Ethnographic observation was used that included both live observations with field notes and also video-recording to capture the design process and its outcomes for detailed review.

Participants were divided into teams of two or three people, and each team was given one kit of materials to use; either:

- paper and pencils,
- card and glue, or
- modelling clay (commonly known as Plasticine).

Participants were instructed to use the allocated materials as detailed above, but beyond this they were not told how to use the materials. In practice, the materials suggested ways of use – for example, no team in the paper and pencil group chose to fold or mould the paper to make a model; as expected, they produced sketches of their design. In everyday design practice, any or all of these materials would be used according to the preferences of the designer and the nature of the design problem. However, in this exercise, participants were restricted to just one kind of material. Thus we were performing something similar to a 'breaching experiment' (Garfinkel, 1967) which deliberately disrupts human activities in order to bring to light aspects that are tacit or taken for granted; although, in the standard breaching experiment the conventions broken are social, whereas here we are disrupting the ability to choose appropriate materials.

Materials

The workshop attendees were a mixture of designers and technologists, artists and architects, psychologists and philosophers. We opted for materials that are accessible for a wide range of people and which did not require any specialized skills. The choice of materials also reflects traditional design practice and covers a range of properties, such as, two-dimensional (2D) versus three-dimensional (3D), manual versus cerebral, malleable versus constrained.

Pencil and paper are of course used extensively throughout the design process for sketching, a process by which the designer works on the design problem by exploring various ideas and experiments with different approaches.

Card and glue is an extension of this – allowing very quick rough '3D sketching' to give ideas some shape for discussion, although card can also be used for more refined models.

The most common material used by product designers to create fast 3D models is 'blue foam', however it was not used in this exercise because of the skills, tools and training required as well as accommodation issues (blue foam is very messy and produces fumes when cut with hot wires). Instead, modelling clay was chosen as it has a long history of use in the 3D design process by many designers 'according to the type of product and company practice' (Bordegoni and Cugini, 2005; Verlinden *et al*, 2001), most notably by the automotive industry (Rekimoto, 1996). Modelling clay provided the tactile 3D element of blue foam without the need to train the participants to use specialist equipment.

Design brief and setting

The brief was to design a hand-held device for producing light that can be turned off and on (see Figure 1). Technical considerations were deliberately kept to a minimum with the intention of encouraging participants to reflect on the device in relation to the human body.

> This exercise looks at the physical design process and its effect on the physicality of the design concept.
> We have split you into 3 groups with the same design brief but with different design materials –one will have pencils and paper; the other paper, card, scissors and glue, and the third, plasticine.
> We will be recording the process and outcomes for later review. At the end of the exercise we will invite each team to talk about their concept, the process they undertook and the pros/cons of the tools used.
> The brief is to design a hand-held device for producing light – it should be able to turn on and off.

Figure 1
Design brief given to participants.

Participants were divided into three groups, one for each material. Each group had two to three teams consisting of two or three people. Figure 2 shows the composition of each group. Each group was given the same design brief but different materials to work with. The teams are labelled team A–H and the participants' names have been anonymized in the transcript fragments presented here.

The participants came from various disciplines including computing, arts, design, sociology, philosophy, human geography

Group	Team	Team members
Card and Glue	A	$A1^{+}$, $A2^{*}$, $A3^{+}$
	B	$B1^{*}$, $B2^{+}$, $B3^{*}$
Plasticine	C	$C1^{*}$, $C2^{*}$, $C3^{+}$
	D	$D1^{+}$, $D2^{+}$
	E	$E1^{+}$, $E2^{+}$, $E3^{+}$
Paper and Pencils	F	$F1^{*}$, $F2^{+}$
	G	$G1^{*}$, $G2^{+}$
	H	$H1^{*}$, $H2^{*}$, $H3^{+}$

Key
* male
\+ female

Figure 2
Allocation of teams to groups.

and architecture. We had eight teams in all; they were given forty minutes to work on the exercise and then invited to present their concepts and comment on the materials they had used. All the teams were based in the same large meeting room but each team worked independently.

Initial observations

Video recordings from three camcorders and photographs were taken during the design exercise of both the individual teams and the final presentations. As there were more teams than video cameras, we chose to rove between teams within a design material group and have a series of shorter recordings showing periods of group interactions rather than continuous end-to-end video of single teams. The trade-off between breadth of coverage of different teams and depth for a single one does not appear to have a simple answer, but in this case the more exploratory nature of the study suggested the former. The following discussions show that the periods of videoing were sufficient to capture significant events in full.

The level of exploration varied significantly between teams. Some teams focused on a single design idea and produced a single prototype, whilst others explored various design ideas and generated a number of prototypes. However, there was no clear relationship between the groups that used particular materials and those being more prolific or more focused in terms of process and output. Indeed in each group there was at least one focused team and one more exploratory team.

Participants came up with a variety of designs, from fairly traditional functional torches, to a child's bedtime cuddly toy that glowed when stroked. Figure 3 shows some of the prototypes that the participants produced during the design exercise. One team spent most of their time in discussion rather than using the materials as a means to explore designs, and defied the instructions that they should not write things down. It was only at the very end that they used the Plasticine to implement an already complete design idea. Another team that only had paper and pencil to work with was the most prolific in terms of the number of design ideas they produced.

Figure 3
Sample of prototypes produced with different materials.

In other cases, the nature of the materials drove the design, so one of the teams working with card ended up producing a cylindrical prototype.

However, it is not possible to identify a direct relationship between the type of material used and the approach each team adopted. Detailed analysis of the video recordings reveals subtle patterns of behaviour including: the way groups move between abstract and concrete discussions, the way different groups either comply with or work against the materials they are given, and the complex interactions between the physicality of materials and group dynamics.

Interactions within a team

We will now examine one team's interactions step by step to provide an understanding of team behaviour and also begin to highlight issues, which we will discuss more thematically in the latter part of the paper.

The team under consideration here is team A, which was supplied with a range of card sizes (A4, A3 and A0 rolls), a glue stick, masking tape and a pair of scissors alongside the design brief. It is interesting to note that there were some remarkable differences between the two teams in the card and glue group, teams A and B, both in terms of the way the members collaborated and in the type of prototypes produced. Also, the participants from both teams employed a lot of gestures during the discussion to demonstrate the ideas they were trying to get across.

The participants in team A are denoted as A1, A2 and A3. The team members spent a substantial amount of time discussing their design concept and exploring various alternatives with the materials they were given. They began with an obvious idea of a torch as a hand-held device that can be switched on and off, but they soon moved away from that concept, as A2 later confirmed in the presentation session.

> A2: We started with the obvious torch, you just press a button to light . . . we thought that was very boring . . .

They narrowed their design focus by thinking of a possible scenario that they may need such a device for.

> A2: . . . we thought about what you need light for, we came up with the very plausible scenario of you wanting to read under your blanket without disturbing or being disturbed.

They went on to explore the shape of the device. A1 starts by naturally rolling the card into a cylindrical shape, but A3 suggests 'what if it can be a handheld device, that is really flat and you can unfold and keep in your pocket'.

The type of light source was the next issue that was discussed.

> A2: Something that illuminates like a keyhole.
> A3: It can be any light, a strobe light . . . an instant strobe light.
> A1: A head torch?
> A3: What about something rechargeable, that's not very heavy?

The team members carried on with their discussion until their ideas began to develop. They proceeded to make prototypes and ended up producing three main prototypes, one based on each team member's design concept. However, there was some degree of collaboration between the team members during the model-making process, as described below. Their aim was to produce a reading light that is inconspicuous and more importantly, does not look like a traditional reading light.

Prototype 1 (Team A)

A2 starts by rolling up a piece of card (A4 size) and taping it with some masking tape to make a tube. A1 is quick to point out the issue with using a straight tube in a tent.

> A1: In a tent, the problem is that with a straight tube light [demonstrates using the glue stick] the angle is wrong, what you want is a torch that bends as an upside down U shape on the top [demonstrates using gestures].

A1 proceeds to make a prototype (see Figure 4) by using another A4 size card (see Figure 4a), she cuts angles off with the pair of scissors (Figure 4b) and uses the masking tape to join the edges (Figure 4c). She cuts open the end to show where the light shines out from (Figure 4d). She later improves the design by adding another piece onto the end (Figure 4e).

During the presentation, A1 describes her thoughts behind her prototype as follows:

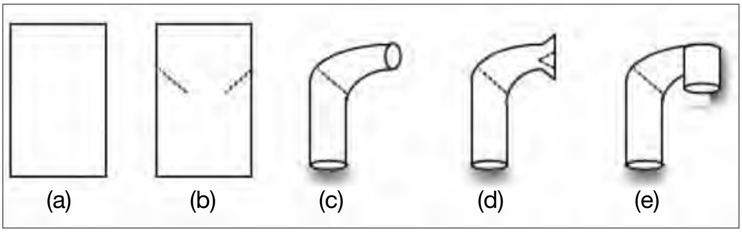


Figure 4
Building stages for prototype 1.

A1: A lamp, the same idea as reading in a tent or reading under the blanket, you are able to hold it in the hand, you want the light to only go on the book, but not on the cover, so you don't get caught by your parents . . . and then it would be good if it's heavy at the bottom so it doesn't, so you don't need to hold it.

Prototype 2 (Team A)

A2 rolls out an A4 size card into a thin tube to produce a reading light that can be attached to the outside of a book, which he demonstrates by folding a piece of card to represent the book (see Figure 5). A3 suggests having a flexible light at the top so one can easily point to different places on the page but A2 remarks that 'a "V" shaped is better for shedding even light across the whole page'. A2 adds 'it is better to stick it outside the card otherwise you can't flip the pages of the book'.

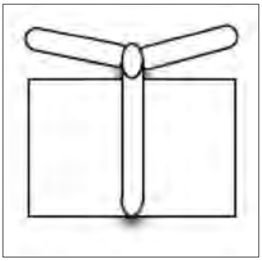


Figure 5
Example prototype 2.

They then engage in a discussion as to where the batteries should be fitted. A2 suggests 'on the flat side (showing the book spine) . . . we need flat batteries'.

A2 later describes his prototype as follows:

A2: A book reading lamp for underneath your blanket, you basically hold it behind your book, and there is a little switch here (imaginary one) . . . you move it up and down, it shines light on your page, you can flip the pages without the light being in your way, and you can hold it in one hand and hold up the blanket with the other hand.

Prototype 3 (Team A)

A3 scrunches a piece of card to demonstrate an 'ergonomic hand-held' shape that can be shone onto things. She engages her team members by suggesting:

A3: Imagine if it is made out of rubbery material (holding the scrunched up paper) that you can squeeze and the best thing is that it doesn't look like a torch, so if you get caught you don't have a torch, you have a stress ball or something!

A1 suggests 'a teddy bear on the wrist' and A3 embraces that idea.

A3: Oh yes a teddy bear, so the light comes out of its eyes … how do you make that … ok, I try to build that one now.

A3 starts by making a pattern for the teddy bear. A1 and A2 join in with their own teddy bear models made from different size cards. A1 cuts out a fairly large flat teddy bear. A2 instead uses the ergonomic shape that A3 made earlier as a mould and adds layers to it, resulting in an amorphous hand-held shape.

A3 however works towards making a medium-size 3D teddy bear model. She makes the 3D body shape from two layers of card cut-outs and sticks the cut-out angles together to produce a 3D effect. She adds some scrunched up card as stuffing between the layers. She then proceeds to make a 3D head shape with some stuffing in between. Before joining the head to the body, she adds some flat arms and secures everything with the masking tape.

Looking at the end result (see Figure 6), A3 remarks 'it is quite a big teddy lamp!' A2 calls it 'iTeddy'. A1 adds a belly button to her flat teddy model and suggests 'this can be pushed, like that (holding the teddy up)'. A3 sticks a button on her 3D teddy model too but A2 remarks 'I don't think we need the button, in a way, you can just squeeze it (demonstrates using the model) to switch it on'. A3 agrees and removes the button and says 'keep it as a conceptual sketch … and how it's gonna look, the light is gonna come out [squeezing the teddy]'.

A3 later explains the thoughts behind her design:

A3: When I was little I used to take my microscope lamp to read, it wasn't easy to switch off. So I thought of something that was hand shaped that you could squeeze … and then we did this one (showing the teddy bear) and then thought it would be cool, if your parents actually catch you and you don't

Figure 6
Finished prototype 3.

Figure 7
Variety of prototypes from Team A.

> have a lamp in your hands, so you have like a teddy bear ... you just press the teddy and the light shines ... we thought a book and it would be a good supplement for a children's magazine.

Figure 7 shows a selection of the prototypes produced by Team A.

Thematic Analysis

After the event, the videos were transcribed by one of the authors who had been present during the study. The video was then informally analysed with the aid of the transcription by another author who had not been present during the design exercise, thus allowing a degree of distance from the data. The aim was to pick up any trends or patterns from the data gathered. The transcription included both the words spoken and descriptions of the artefacts being prototyped, sketching, gestures and so on. The dialogue and descriptions in the previous sections are excerpts from this transcript.

The transcripts were then subjected to a systematic in-depth iterative analysis. The analysis began with a small number of known themes and issues, principally physicality itself and its influence on creativity in the designs produced. This was used to pump prime iterative passes over the data using dialectic re-coding (Dix, 2008). This involves coding the data according to the known categories, but not doing so in order to 'gather evidence' for pre-judged themes, but instead subjecting them to critique. We looked for two main kinds of tensions (the dialectic) between the themes and the data:

(i) gaps in coding – apparently important events or statements that did not fit within the thematic scheme;
(ii) inadequate coding – where there is a way of coding the data, but it appears to be incomplete in its description of the data.

The first leads to the identification of new themes, many of which had not been considered. The second leads to refinement or modification of existing themes.

This process combines an inductive data-driven qualitative analysis with an explicit recognition that we had pre-existing expectations and intentions when approaching the data gathering and data analysis. This is similar to the Straussian School of grounded theory (Strauss and Corbin, 1990) and indeed dialectic re-coding was first proposed as a way of validating the results of grounded theory or other qualitative analysis methods.

Even within Team A, we can identify a wide variety of behaviours: from designs driven by the physical properties of the materials, for instance when A2 rolls up a piece card in prototype 1 and 'finds' a classic cylindrical torch shape, to more abstract discussions of properties:

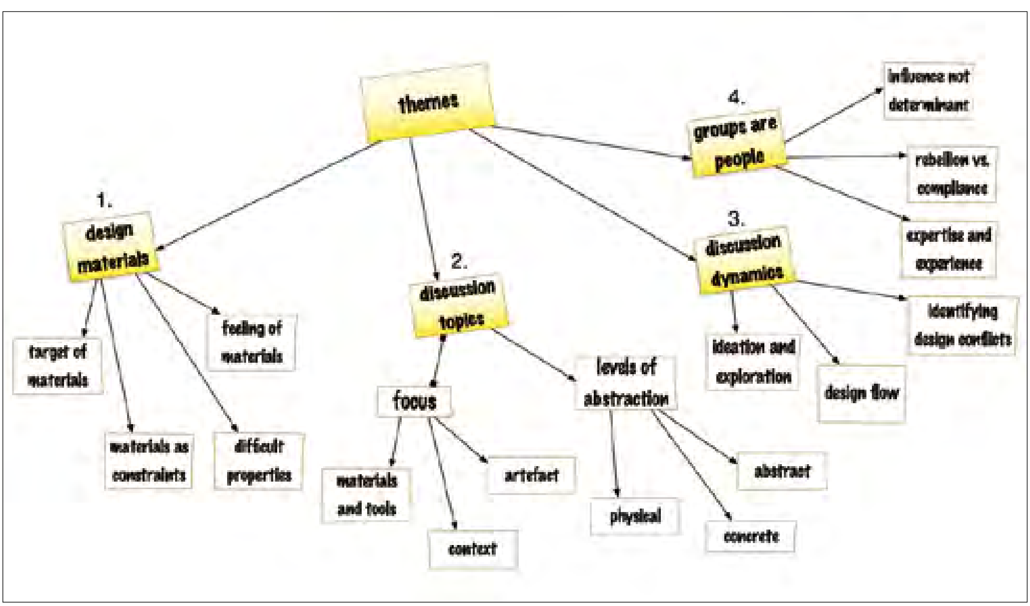
A3: What about something rechargeable…

There were also some underlying tendencies, for example, the groups with paper and pencil tended to produce more fragments of ideas, but not necessarily more finished design concepts. However, the picture is typically more complex, and the themes have helped us unpick the rich interplay between materials, design brief, team make-up and dynamics.

We have categorized the themes into four main classes (see Figure 8):

1. Those relating specifically to what the design materials give to the teams;
2. Those concerned with the topic of the design discussions;
3. Those relating to the flow or dynamics of design discussion; and
4. Those relating to the personal and interpersonal factors within the group.

Figure 8
Breakdown of themes.



We will briefly discuss the first two of these themes to draw out issues related to the physical properties of the materials during the early design phase.

Theme 1: Design Materials

Target of materials: prototype/design/product

Some of the materials given to the groups were intended to be used for prototyping: card and Plasticine for example, whilst others, such as paper and pencil although used during the design process are not evident in the prototype itself (except insofar as the design sketches are on paper).

None of the teams given only paper and pencil used these materials to create a paper model by folding, tearing, and so forth, despite having ample paper to draw on and use for prototyping. It can be identified that once participants regarded the paper as an object for drawing on (design material), it became unlikely that they would see it as a material for construction. This convention was seen despite many other forms of challenging behaviour, thus suggesting that it is a hard mindset to change.

Occasionally, the discussion turned to the actual materials that would be used on the product assuming the design were realized, for example,

> C1: . . . and LumaTed is made from luminous material – 'Philips Lumalight', with a sort of translucent material, that can light up in different colours, and you interact with it by stroking, so it will have a set of capacitor sensors that will allow you to stoke it in different ways and the more you stroke it the brighter it gets . . .

This group even went on to give LumaTed a price. However, few groups explicitly differentiated the prototype material from the production material, probably exacerbating some teams' tendency to be 'trapped' by the materials (see discussion below).

The materials included scissors, masking tape and so on, and were sometimes used in unexpected ways. For example, team B considered using a roll of masking tape to draw smaller circles, and the scissors as a compass to draw larger ones using the point as a pivot. Furthermore, materials were often drawn in from the environment in addition to those supplied. Team D used a water bottle extensively both to inspire their design and eventually as part of their prototype (that is both as a design material and a prototype material). Another team rolled Plasticine on the rough walls in an attempt to produce textured surfaces.

Ethnographies of group activities have demonstrated the importance of shared artefacts in coordinating actions (Heath and Luff, 1992). Likewise, in our study the design materials are used to manage sharing and provide shared focus. For example, in Team F, members F1 and F2 started sketching ideas on separate sheets of paper, but rapidly switched to using the same sheet of paper, thus using the paper as a shared artefact to maintain collaboration and generate design ideas.

Materials as constraints

Constraints can sometimes be seen as a bad thing that limits or holds back development. However, psychological research on creativity and problem solving has often found that constraints can inspire creative designs (Ormerod *et al*, 2002) partly because they focus the design and partly because they sometimes prevent the ‘obvious’ solutions.

As we saw earlier during the production of Prototype 3 (Figure 6), A3 did not let the card, an essentially 2D material, hinder her design for a 3D teddy bear model. In fact her ability to produce such a refined model most likely reflects her background in fashion design.

However, most of the other participants often referred to the fact that they would have preferred one of the other materials, sometimes during specific parts of the design process. Some wanted pens to sketch with or another set of materials to use, for instance with Team A:

> A2: Yes, I mean, I think we were lacking some pen to scribble.
> A3: Or Plasticine!

And Team B:

> B2: I think to sketch out an initial design probably would’ve been handy, or a way to, sort of, come up with some ideas.

Similarly, Team H were frustrated that they have not been given clay to try out their ideas:

> H2: I think ... the sketching, or the pencil and paper were ok for the initial communication of ideas and ... summarising what we thought was right...
> H3: Yeah, it would’ve been good to have something that we could actually mould and actually get more feeling about the actual prototype we came up.

Noticeably there was a greater tendency to ask for clay than card. This may reflect its greater malleability, for example Team A admitted during their presentation that they found the card ‘difficult to bend to the shapes that our minds had formed in our head’ and Team F ‘...tried to look at some sort of more organic bioforms, shapes, but paper is not a very good medium for doing that...’. The popularity of the clay may well also be because it was regarded as being more ‘fun’.

Teams also responded more subtly to their materials: the majority of card-based prototypes were formed from cylinders and other rollable shapes. However, this material did not totally determine the design, as we saw Prototype 3 (Figure 6) included scrunching up the paper, making it in effect more like the clay.

Difficult properties

Some properties that were mentioned during design discussions were difficult to replicate with physical prototypes. This included the weight of objects (hard with card), softness (hard with Plasticine), the light itself.

> F1: How do we produce light? . . . is a spark a light?'

We note that only the teams using paper and pencil discussed energy and light at length. This is perhaps related to the fact that they engaged in more abstract discussions and that whilst weight, energy and light are very physical in one sense, they are also somewhat ethereal properties. But even this was not a universal rule as Team E used a cone of paper to 'simulate light'.

Feeling of materials

Whilst the materials did not determine team behaviour, the fact that Plasticine was a child's toy certainly seemed to influence the teams' attitudes toward material properties. The teams using Plasticine appeared to operate in a more playful and sometimes wacky manner. For example, at one point team E produced a Petri dish while Team C spent a period discussing ideas, but all the while each holding a piece of Plasticine, kneading and playing with it, but not using it to make anything. Even during the discussion stage, C1 starts playing with the Plasticine and says 'Oh yeah, well we really enjoyed the Plasticine! Yeah, that was fun'.

In contrast, card suggested more formal/serious designs:

> B3: In terms of our process it was very much orientated to what we thought we could do with the materials we'd been given . . . we feel that it was good for making something that was solid, if you drop it, it probably won't break, but more than that it's not very expressive . . .

Theme 2: Discussion Topics

The topics that were discussed by the teams during the design exercise were based on different levels of abstractions (from physical, concrete to abstract) and focus (on the artefact, the context and the materials). Some teams spent more time in one or other kind of discussion, but also each team moved between kinds, at one moment discussing concrete design ideas, at another, a more abstract discussion of requirements.

Level of abstraction: physical/concrete/abstract

Some of the discussion focused around the physical nature of the materials and models that were in their hands. For example, team B focused on making a torch that was as realistic as possible:

> B3: . . . most of our process was about making a model that looks relatively realistic, or at least as realistic as we could get.

They also used objects in the environment (such as the water bottle mentioned previously) to augment their design or to demonstrate or stimulate ideas:

> G1: So the idea was a watch with light . . . so the concept was this watch (showing his wristwatch), putting some lamps inside the clock/the watch, by sensors, touching it you can make it work.

And even their own bodies:

> H2: yeah sure . . . you are just cupping the light . . . [demonstrates two different ways of cupping using his hands] . . . I think it would be great . . . for warming light . . . [rubbing hands together] . . . (Note: team H was in the paper and pencil group, so had no obvious prototyping material to create this sort of physical focus for their discussion.)

At other times, the discussion was still quite specific, talking about a particular design or scenario of use, but without having it physically to hand. For example, team C were refining the shape of their teddy bear:

> C1: Yeah to show . . . we have a version . . . when the child grabs, cuddles it, it will come on, that's one situation, or strokes it . . . we have yours where it's at the end of the bed and we have to look for it, so we have to ask it, call its name and it'll come on.

Finally, there were times when the discussion was at a more abstract level discussing general ideas, properties or dimensions. For example team H discussed ideas of 'discrete feedback' and needing some form of 'discrete interaction' and Team B considered the 'primitives' afforded by their material (card):

> B3: We've got a circle (holding the masking tape) so we can use that to create a precise circle, we can mark it with that but we can score the card, so those are the primitives . . .

There was a tendency for the paper and pencil group to have more abstract discussions, but this is far from being their preserve and many teams engaged in some form of more abstract discussion. What was evident was that at the point at which discussions became more abstract the teams with prototyping materials ‘stepped back’ from the materials … and one Plasticine team even ‘cheated’ and used paper and pen!

Despite this ‘stepping back’, this is not to suggest that these different levels of abstraction are independent discussions. On the contrary, there is a constant interplay where more abstract discussions lead back to concrete design suggestions:

> F1: … produces light and there are implications with that … it has a battery, it has a bulb – that’s the normal way to produce light, although they could have one of those, er, wind up ones…

Or lead to physical design solutions:

> H3: I guess if you’re looking for discrete … what you need is some sort of discrete interaction like clapping…

And even physical or concrete considerations prompt generalizations.

> H2: [Sketching on paper] I think if you look at the fire, there’s a couple of things that you can read from the physicality of the fire. If you place more logs on the fire, you can see how long the fire might eventually burn … and if the fire dims, you see the flames going down, you see that you have to put more on the fire.

Focus: artefact/context/materials and tools

It is evident that the discussion topics were concerned with the artefacts that were being designed:

> F2: We started with the obvious torch.

Sometimes on the context in which the artefact would be used:

> A2: We thought about what you need light for, we came up with the very plausible scenario of you wanting to read under your blanket without disturbing or being disturbed.

And sometimes on the materials, tools or process of design itself:

> E2: We used all kinds of tools, we split a pen apart and used the top and cut with this and used round things as forms and

used the paper to simulate light and we borrowed some of the green Plasticine from the other group.

Each of these can be discussed at each level of abstraction as exemplified in Figure 9. We have seen several instances of these, such as in team A's prototype 2, they not only made the reading light, but they also used a folded piece of card to simulate a book: a physical model of the context.

Figure 9
Level of abstraction versus focus of discussion topics.

Level of Abstraction \ Focus	artefact	context	materials/ tools
physical	holding torch	making book	rolling the card
concrete	mentions torch	scenario	scissors as compass
abstract	need to be rigid	requirements	'primitives'

Most of those teams with physical materials apt for prototyping spent a considerable amount of time manipulating the physical artefacts. However, as previously noted, the abstract parts of this space are not the preserve of the paper and pencil group only, indeed team C, in the Plasticine group, at one point raised the following:

> C1: And we decided to focus it on children. And it seemed reasonable in that case to try to make it into some kind of night light, something that will help children when they're feeling frightened at night. And so from that we got a set of properties that we thought we would want to express through this, we wanted to help children feel safe and secure at night, it would be something that would be easy to interact with when they're kind of in that semi-wakeful state, something that would be soft, or warm, smooth, stable, robust ... and out of all that came LumaTed.

The above excerpt illustrates the flow of the discussion which starts off in abstract context 'focus on children', moves to concrete artefact 'night light', then back to abstract context 'feeling frightened at night', to abstract artefact 'set of properties', and eventually back to more refined concrete artefact 'LumaTed'.

Discussion and Conclusion

Our study has demonstrated that a simple design brief combined with low-fidelity materials can generate a wealth of information, thus reaffirming the importance of producing low-fidelity prototypes at an early stage in the design. As Rettig (1994) remarks, 'lo-fi prototyping requires little more in the way of implementation skills other than the ones learned in kindergarten'. The participants from our study were not all 'natural designers'; they came from different disciplines and as the materials they were given did not require detailed skills, they could focus entirely on the features of the design itself.

Our research focus was to explore the role that the physicality of materials plays in the design process. The results do not support simplistic conclusions such as 'physicality promotes creativity', or even the opposite. Our in-depth thematic analysis however reveals dimensions along which general trends can be seen. Although we have focused on the materials and discussion topics, we cannot ignore the effects that the flow or dynamics of the discussion, and the personal and interpersonal factors within a group have on the design process.

Individuals within groups and the way they worked together, would often mean they defied the restrictions or paths suggested by their materials – including rebelling completely, as with the Plasticine team who, against the rules, used paper and pencil. This ability to move against the natural affordance of physical materials was dependent on the characters of individuals and/or teams.

Here are some of the key themes from our analysis:

Physicality and teams – Paper and pencil were firmly regarded as sketching or drawing material, unlike card and clay, which were treated as modelling materials. Like Kingsley *et al*'s study (2005), we found the modelling teams had more fun, felt happier with their end design and more committed to the goals of the group, however, we did not specifically measure for these. Also, as the actual building process started, the level of interaction between members of the modelling team members did slow down, especially with the teams that produced more than one prototype as each one started making on their own design. Similarly, we found that with the paper and pencil teams, the drawing was mainly done by one person, allowing or compelling the team members to continue building their relationship through gaze rather than focusing on their own activity.

Physicality and user experience – We cannot overlook the fact that the way that materials were utilized was partially a consequence of preconceptions brought into play by the backgrounds of the participants. Within a single group, card is treated as if it were clay (crumpled to conform with the shape of a clenched hand), as an essentially 2D material (forming a bear as a cut out) and finally as a textile (forming a 3D teddy). So while the materials supplied may influence the output, it was also clear that the experience the user brought to the table partially influenced the design activity.

Physicality and level of abstraction – Typically the tendency for the teams with paper and pencil was to engage in more abstract discussion as they had no physical focus, thus contributing to the greater number of (often fragmentary) design ideas. When some of the teams in card and modelling clay groups did discuss abstract properties, they had to step back from the material and some even cheated by writing on paper. On the whole, the card and modelling clay groups were more inclined to discuss the physical artefact they were designing, the context in which it would be put to use and how to work with the materials.

For practising designers, but even more so for students, the findings from our research do prompt questions as to how to maximize the benefits of specific physical materials in prompting new ideas, whilst also at appropriate moments during design activity, to 'escape' the practical and cognitive limitations they create. We do not answer this question here, nor can we make any sweeping statements as to the impact of our findings on the way we might design things in the future, but we believe that the rich understanding of the design process we have produced is a step towards this.

Based on Ulrich and Eppinger's (2003) classification of prototypes, in our design exercise, card and modelling clay were used to produce physical prototypes (tangible artefacts that approximate the product) while paper generated more abstract prototypes (where the products are more analysed rather than built). But they were all low-fidelity prototyping materials that are versatile and accessible; as a result, they can be very useful for early exploration of design concepts in a short space of time. They allow the team to try out lots more ideas and are driven by the experience and behaviour that the people bring to the table. In terms of purpose for building prototypes, whilst Ulrich and Eppinger begin with 'Learning', this is expressed in terms of learning about the suitability or potential of an already part-formulated design concept. In contrast our groups had a playful and open brief and their purpose was largely exploratory, closer to Gedenryd's (1998) notion of design as inquiry. Also while Ulrich and Eppinger suggest that only physical prototypes can be 'comprehensive', the importance of context of use highlighted in the last section, questions this comprehensiveness, which is perhaps more about the final artefact in isolation, not the artefact in context.

Theoretical understanding of physical artefacts is focused on objects that concretely achieve physical goals: for Heidegger (1962) the way a hammer is 'ready to hand' in the act of joining wood with nails, or for Gibson, the way a rock if of suitable size 'affords' sitting upon. In the case of objects in a 'natural' (pre-technological) environment, Gibson argues that if we are well adapted to the environment, then our perceptions are tuned so that the affordances are immediately perceived; we are creatures tuned for action. However, as soon as we consider technological objects, things become more complex. Even turning a door handle needs to be considered as a sequentially unfolding chain of learnt associations and skills, as well as more immediate visual and haptic perceptions (Gaver, 1991). Similarly, 'ready to hand', while frequently misquoted, is not a matter of 'walk up and use' but is the product of culture and skill; indeed Ilyenkov (1977) regarded ideals (ideas) as embodied relationally in the activities of creating and using artefacts.

In this paper, we have looked at materials in design – that is physical objects that are for essentially cognitive tasks. What a material 'affords' under such circumstances is even more finely dependent on the past knowledge and skills of those using them (for example, fashion designer versus sculptor); and yet the material is not entirely open, without influence, like a piece of wood being carved, it has a grain, a set of uses that are easier than others, that fall more readily to hand or mind. Building an adequate practical and theoretical understanding of such a nuanced and context sensitive area is no easy task, and one we have by no means accomplished, but is, we believe, a valuable goal.

Philosophers of embodiment such as Clark (1998) and Gallagher (2005), and Gedenryd's (1998) application of this strand of thinking to design go some way towards this. However, some of our results, notably the way the teams turned to discussion or paper and pencil for more abstract 'stepping back' or Team A's 'shapes that our minds had formed in our head', seems to contradict Gedenryd's claim that 'designers go out of their way to avoid intramental thinking'. This may be because our participants were academics, used to talking (!), and maybe had a rich repertoire of concrete examples that could be drawn upon mentally. In contrast, Gedenryd's focal transcripts concern tutor–novice dialogue, where the tutor is possibly being concrete for the benefit of the novice. However, looking afresh at even these transcripts (Gedenryd, drawn from Schön, 1983) reveals a more mixed picture, as the pivotal action of the tutor, Quist, is to notice that the novice, Petra, has become stuck in detailed design and so needs to step back to reformulate the problem. The outcome of this stepping back is presented concretely in sketches representing an alternative approach, but the stepping back itself is conveyed in words.

Looking forward, future empirical studies could focus on practising product designers, observing a longer exercise over which participants could get more involved in the process and tools they are using. Rather than limiting them to a single prototyping material, a combination could be offered at different stages of design to observe which are chosen, how they are used, and indeed when they are not used at all. More interventionist experiments could control the order the materials are made available to groups, maybe presenting materials deliberately in the 'wrong' order, for example, first modelling clay followed by paper and pencils, then card and glue.

In summary, the data resists simplistic assumptions. Choice of material clearly has an impact, but the individual skills and background of the participants are equally important. Materials can constrain people, but they can also inspire creative design, hence the need for physical modelling to be part of the design curriculum and students' awareness of and willingness to use these techniques need to be cultivated in an age of CAD. But equally, less tangible expression in discussion and plain old lists seems to allow breadth of exploration. A full theoretic approach needs to take this richness into account, not just the embodied interplay between mind and environment, but also recognizing that parsimony (Clark, 1998) cuts two ways with the balance between mental/verbal and manual/physical 'cognition' dynamically adapting to suit the situation.

References

Avrahami, D. and Hudson, S. E. (2002). 'Forming interactivity: A tool for rapid prototyping of physical interactive products'. In Verplank, B., Sutcliffe, A., Mackay, W., Amowitz, J. and Gaver W. (eds), *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS'02*. New York: ACM Press, pp. 141–146.

Baxter, M. (2002). *Product Design: A Practical Guide to Systematic Methods of New Product Development* . Cheltenham: Nelson Thornes (Publishers).

Bordegoni, M. and Cugini, U. (2005). 'Design products with your hands'. In *Virtual Concept 2005* [Conference]. Biarritz, France, 8–10 November 2005.

Branham, R. (2000). 'Given the radically changing work environment and new worldviews, what kinds of new "tools" do designers need to survive and successfully deal with tomorrow's design problems?' In *IDSA 2000 National Education Conference*. Louisiana, USA, 18–20 September 2000.

Buchenau, M. and Suri, J. F. (2000). 'Experience prototyping'. In Boyarski, D. and Kellogg, W. A. (eds), *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*. New York: ACM, pp. 424–433.

Clark, A. (1998). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Dant, T. (1999). *Material Culture in the Social World: Values, Activities, Lifestyles*. Buckingham: Open University Press.

Dix, A. (2008). 'Theoretical analysis and theory creation'. In Cairns, P. and Cox, A. (eds), *Research Methods for Human–Computer Interaction*. Cambridge: Cambridge University Press, pp. 175–195.

Dix, A., Ghazali, M., Gill, S., Hare, J. and Ramduny-Ellis, D. (2009). 'Physigrams: Modelling devices for natural interaction'. *Formal Aspects of Computing* [online]. SpringerLink paper. Available at: www.doi: 10.1007/s00165-008-0099-y [accessed 27 July 2009].

Evans, M. and Sommerville, S. (2007). 'Seeing is believing: The challenge of product semantics in the curriculum'. In *Shaping the*

Future? 9th International Conference on Engineering and Product Design Education. Northumbria University, Newcastle Upon Tyne, UK, 13–14 September 2007.
Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.
Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
Gaver, W. (1991). 'Technology affordances'. In Robertson, S. P., Olson, G. M. and Olson, J. S. (eds), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching through Technology*. New Orleans, Louisiana, USA, 27 April–2 May 1991. New York: ACM Press, pp. 79–84.
Gedenryd, H. (1998). *How Designers Work – Making Sense of Authentic Cognitive Activities* [online]. PhD thesis. Lund University, Sweden. Available at: http://www.lucs.lu.se/Henrik.Gedenryd/HowDesignersWork/ [accessed 27 July 2009].
Ghazali, M. and Dix, A. (2005). 'Knowledge of today for the design of tomorrow'. In *Proceedings of the 2nd International Design and Engagibility Conference (IDEC 2005)*. Edinburgh, UK, 6 September 2005.
Gibson, J. J. (1979). 'The theory of affordances'. In Shaw, R. and Bransford, J. (eds), *Perceiving, Acting and Knowing: Towards an Ecological Psychology.* Hillsdale, NJ: Erlbaum, pp. 67–82.
Gill, S., Johnson, P., Dale, J., Loudon, G., Hewett, B. and Barham, G. (2005a). 'The traditional design process versus a new design methodology: A comparative case study of a rapidly designed information appliance'. In *Home Oriented Informatics & Telematics Conference*. University of York, UK, 13–15 April 2005.
Gill, S., Loudon, G., Hewett, B. and Barham, G. (2005b). 'How to Design and prototype an information appliance in 24 hours – integrating product & interface design processes'. In *6th International Conference on Computer Aided Industrial Design and Concept Design Conference.* University of Delft, The Netherlands, 29 May–1 June 2005.
Gill, S., Walker, D., Loudon, G., Dix, A., Woolley, A., Ramduny-Ellis, D. and Hare, J. (2008). 'Rapid development of tangible appliances: Achieving the fidelity/time balance'. In Hornecker, E., Schmidt, A. and Ullmer, B. (eds), *International Journal of Arts and Technology (IJART): Special issue on Tangible and Embedded Interaction*, 1(3/4), 309–331.
Greenberg, S. and Fitchett, C. (2001). 'Phidgets: Easy development of physical interfaces through physical widgets'. In Marks, J. and Mynatt, E. D. (eds), *Proceedings of the 14th Annual ACM Symposium on User interface Software and Technology, UIST '01*. Orlando, Florida, 11–14 November 2001. New York: ACM Press, pp. 209–218.
Hartmann, B., Klemmer, S. R., Bernstein, M. Abdulla, L. Burr, B., Robinson-Mosher, A. and Gee, J. (2006). 'Reflective physical

prototyping through integrated design, test, and analysis'. *Proceedings of UIST 2006.* New York: ACM Press, pp. 299–308.

Heath, C. and Luff, P. (1992). 'Collaboration and control: Crisis management and multimedia technology in London underground line control rooms'. *Computer Supported Cooperative Work*, 1, 69–94.

Heidegger, M. (1962). *Being and Time*. London: SCM Press.

Heskett, J. (1980). *Industrial Design.* London: Thames & Hudson.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Ilyenkov, E. (1977). 'The concept of the ideal'. In *Philosophy in the U.S.S.R: Problems of Dialectical Materialism*. Moscow: Progress Publishers. Available at: http://www.marxists.org/archive/ilyenkov/works/ideal/ideal.htm. Transcribed by A. Blunden.

Kingsley, S., Baxter, S. and Inns, T. (2005). 'Educating the designer for team working: An experiment on the effects of prototyping on teams'. In Rodgers, P., Brodhurst, L. and Hepburn, D. (eds), *Crossing Design Boundaries*. London: Taylor & Francis, pp. 243–247.

Landay, J. and Myers, B. (1995). 'Interactive sketching for the early stages of user interface design'. In Katz, I. R., Mack, R., Marks, L., Rosson, M. B. and Nielsen, J. (eds), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press/Addison-Wesley, pp. 43–50.

Norman, D. (1998a). *The Design of Everyday Things*. London: MIT Press.

Norman, D. A. (1998b). *The Invisible Computer: Why Good Products Can Fail, the PERSONAL COMPUTER Is So Complex and Information Appliances Are the Solution*. London, Cambridge, MA: MIT Press, p. 189.

Ormerod, T., MacGregor, J. and Chronicle, E. (2002). 'Dynamics and constraints in insight problem solving'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 791–799.

Pevsner, N. (1991). *Pioneers of Modern Design: From William Morris to Walter Gropius*. London: Penguin Books.

Phidgets Inc. (2008). *Phidgets* [online]. Available at: *http://www.phidgets.com/* [accessed 27 July 2009]

Pierce, J. S., Steams, B. C., and Pausch, R. (1999). 'Voodoo dolls: Seamless interaction at multiple scales in virtual environments'. In Rossignac, J., Hodgins, J. and Foley, J. (eds), *Proceedings of the 1999 Symposium on Interactive 3D Graphics*. New York: ACM Press, pp. 141–145.

Potter, N. (1969). *What Is a Designer: Things, Places, Messages*. London: Studio Vista.

Raizman, D. (2004). *A History of Modern Design: Graphics and Products since the Industrial Revolution*. London: Laurence King Publishing.

Rakers, G. (2001). 'Interaction design process'. In Baumann, K. and Thomas, B. (eds), *User Interface Design for Electronic Appliances*. New York: CRC Press, pp. 29–47.

Ramduny-Ellis, D., Dix, A., Hare, J. and Gill, S. (eds) (2007). *Physicality 2007: Proceedings of the Second International Workshop on Physicality.* Cardiff: UWIC Press. Lancaster University, UK, 2–3 September 2007. Available at: http://www.physicality.org/Physicality_2007/Entries/2008/3/8_Physicality_2007_proceedings_files/Physicality2007Proceedings.pdf.

Rekimoto, J. (1996). 'Transvision: A hand-held augmented reality system for collaborative design'. *Proceedings of Virtual Systems and MultiMedia (VSMM '96).* International Society on Virtual Systems and MultiMedia (VSMM). Gifu, Japan, 18–20 September 1996, pp. 85–90.

Rettig, M. (1994). 'Prototyping for tiny fingers'. *Communications of the ACM*, 37(4), April 1994, 21–27.

Schön, D. (1983). *The Reflective Practitioner: How Professionals Think in Action*. London: Temple Smith.

Snyder, C. (2003). *Paper Prototyping.* San Francisco, CA: Morgan Kaufmann.

Strauss, A. and Corbin, J. (1990). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. London: Sage Publications.

Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human–Machine Communication*. New York: Cambridge University Press.

Ulrich, K. T. and Eppinger, S. D. (2003). *Product Design and Development*. New York: McGraw-Hill.

Verlinden, J. C., Wiegers, T., Vogelaar, H., Horvath, I., Vergeest, J. S. M. (2001). 'Exploring conceptual design using a clay-based Wizard of Oz technique'. In Jahannsen, G. (ed.), *8th IFAC/IFIP/FORS/IEA Symposium on Analysis, Design and Evaluation of Human–Machine Systems.* Kassel, 18–20 September 2001, pp. 211–216.

Vihma, S. (1995). *Products as Representations: a Semiotic and Aesthetic Study of Design Products*. Helsinki: University of Arts & Design, Helsinki.

Villar, N., Lindsay, A. and Gellersen, H. (2005). 'Pin & play & perform: A rearrangeable interface for musical composition and performance'. *NIME '05: Proceedings of the 2005 conference on New interfaces for musical expression*. Vancouver, Canada, 26–28 May 2005, pp. 188–191.

Wheeler, M. (2005). *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.

Woodham, J. (1997). *Twentieth-century Design*. Oxford: Oxford University Press.

Woodward, I. (2007). *Understanding Material Culture*. London: Sage Publications.

Biographies

Devina Ramduny-Ellis is a Research Associate in the Computing Department at Lancaster University in the UK. She has worked on several projects covering areas such as computing infrastructure, user interfaces, ubiquitous computing, dependability, and decision management. Her research interests are diverse: devising methods for work analysis and interaction properties, modelling techniques, employing ethnographic and qualitative techniques to inform novel designs, and physicality and digitality. She has been actively involved in starting up a multi-disciplinary community in physicality through the Physicality 2006 and Physicality 2007 International Workshops and co-editing a journal special issue on Physicality and Interaction. She has over 25 publications and is managing the 'DEPtH: Designing for Physicality' AHRC/EPSRC funded research project as part of the 'Designing for the 21st Century' Initiative.

Jo Hare graduated in 2004 from De Montfort University, Leicester, with a BSc Honours degree in Industrial Design, during which she completed a placement year working within BMW's Design and Development building in Munich. After graduating, Jo worked for Exmoor Plastics, a medical device manufacturing company in Taunton, in conjunction with The National Centre for Product Design & Development Research (PDR) as part of the Knowledge Transfer Partnership (KTP) scheme. The KTP project focused on increasing the company's productivity by decreasing product development time. New procedures were successfully implemented that defined and selected new project proposals. Jo is currently working within the Programme for Advanced Interactive Prototype Research (PAIPR) research group on 'DEPtH: Designing for Physicality', an AHRC/EPSRC funded project under the 'Designing for the 21st Century' programme.

Alan Dix is a Professor in the Department of Computing at Lancaster University in the UK. He was a mathematician by training, and mathematics is still his first love, but he has worked in Human–Computer Interaction since 1984, has published over 300 articles and is author of a key textbook in the area. He has worked in several universities, agricultural engineering research, local government and hi-tech start-ups. His interests are eclectic: formalization and design, physicality and digitality, the economics of information, structure and creativity and the modelling of dreams. Recently he and a colleague have developed technology for autonomous pixels that can be configured and turn any surface or space into a two- or three-dimensional display.

Martyn Evans is a Senior Lecturer in Design at Lancaster University. As a trained product designer, his research interests explore the approaches designers use to consider the future, in particular the ability of designers to envision potential social, cultural, technological and economic futures. He was co-investigator upon ‘Design 2020’, an AHRC/EPSRC-funded research project looking at potential futures for the UK design industry. He has secured and supervised a number of government-funded knowledge transfer initiatives within the area of design and new product development. With extensive experience of leading undergraduate and postgraduate design curricula, he is currently Course Director for MA Design: Management and Policy and external examiner at a number of UK institutions.

Steve Gill is a product designer with 16 years experience in industry and academia. He is Director of Research of the Cardiff School of Art & Design and Director of the Programme for Advanced Interactive Prototype Research (PAIPR) within The National Centre for Product Design & Development Research (PDR). He has designed or product managed around 50 products to market and has published 30 academic journal and conference papers. Steve has a range of research interests related to product design and development including the rapid design development of information appliances, the role of physicality in the design process and the implications for product design development of testing early information appliance prototypes within their context of use. He works closely with academic partners such as the University of Lancaster and those in blue chip industry such as Sony-Ericsson.

Addresses for Correspondence

Dr Devina Ramduny-Ellis, Computing Department, InfoLab21, Lancaster University, Lancaster LA1 4WA, UK.
Tel: +44 (0) 1524 510501
Email: devina@physicality.org

Jo Hare, Cardiff School of Art & Design/National Centre for Product Design & Development Research (PDR), University of Wales Institute (UWIC), Western Avenue, Cardiff CF5 2YB, UK.
Tel: +44 (0) 29 2041 7009
Email: juhare-pdr@uwic.ac.uk

Professor Alan Dix, Computing Department, InfoLab21, Lancaster University, Lancaster LA1 4WA, UK.
Tel: +44 (0) 1524 510319
Email: alan@hcibook.com

Martyn Evans, ImaginationLancaster, The Roundhouse, Lancaster University, Lancaster LA1 4YW, UK.
Tel: +44 (0) 1524 593435
Email: m.evans@lancaster.ac.uk

Dr Steve Gill, Cardiff School of Art & Design/National Centre for Product Design & Development Research (PDR), University of Wales Institute (UWIC), Western Avenue, Cardiff CF5 2YB, UK.
Tel: +44 (0) 29 2041 6732
Email: sjgill@uwic.ac.uk

Acknowledgements
We would like to thank all participants of the *Physicality 2007* workshop for their contribution and the invaluable insights they provided us by taking part in the design session. This work has been supported by the AHRC/EPSRC-funded project 'DEPtH: Designing for Physicality' (http://www.physicality.org/) as part of the 'Designing for the 21st Century' research initiative (http://www.design21.dundee.ac.uk/). Thanks also to the anonymous referees for their critique and for pointing us to fresh and valuable literature.

Rapid Development of Tangible Interactive Appliances: Achieving the Fidelity / Time balance

Steve Gill, Darren Walker, Gareth Loudon, Alan Dix, Alex Woolley, Devina Ramduny-Ellis and Jo Hare

Steve Gill (Corresponding Author)

National Centre for Product Design & Development Research, UWIC

Western Avenue

Cardiff

CF5 2YB

UK

Phone: +44 (0)29 2041 6732

E mail: sjgill@uwic.ac.uk

STEVE GILL is a product designer with 15 years experience in industry and academia. He is a Principal Lecturer at the *University of Wales Institute Cardiff* and Director of the *Programme for Advanced Interactive Prototype Research* (PAIPR) within The *National Centre for Product Design & Development Research* (PDR). He has designed or product managed around 50 products to market and has published 20 academic journal and conference papers. Steve has a range of research interests related to product design and development including the rapid design development of information appliances, the role of physicality in the design process and the role of *Cradle to Cradle* theories in product design & development.

He works closely with academic partners, particularly *Lancaster University* and those in blue chip industry such as *Sony-Ericsson* and *Samsung*. He is currently engaged in a major project with partners at Lancaster University investigating the nature of physical interaction and its effects on design with the help of an Arts and Humanities Research Council grant.

Darren Walker

Cardiff School f Health Sciences, UWIC

Western Avenue

Cardiff

CF5 2YB

UK

Phone: +44 (0)29 2041 7227

E mail: dwalker@uwic.ac.uk

DARREN WALKER is a senior lecturer in Psychology at the *University Of Wales, Institute, Cardiff*. He completed his PhD in Cognitive Psychology from the University of Reading, looking at the role of reminders. Darren has a range of research interests, including the role of music and emotion in cognition, the application of psychological theories to the design of cognitive artefacts and issues surrounding ageing and cognitive performance. Interspersed with academia, Darren has held industrial posts in the Aviation sector. He collaborates with colleagues at the Universities of Exeter, Leeds and Trinity College, Dublin. His work has been disseminated in journals along with national and international conferences.

A.

Gareth Loudon

National Centre for Product Design & Development Research, UWIC

Western Avenue

Cardiff

CF5 2YB

UK

Phone: +44 (0)29 2041 6661

E mail: gloudon@uwic.ac.uk

GARETH LOUDON is a Senior Lecturer in Product Design, Programme Director of MSc Advanced Product Design and a Director of the Innovation Forum Ltd. (which offers training, coaching and consultancy services to help companies create new product and service strategies). He has over 18 years experience in academic and industrial research and has taken several research ideas all the way through to commercial products for large companies such as Apple Computer and Ericsson.

Gareth has several patents to his name and over 35 publications. He has also won many awards including Best Software Product Award at COMDEX Asia, and for his concept design work from IDSA/Business Week. Gareth graduated with a 1st Class Honours degree in Electrical and Electronic Engineering in 1988 and a Ph.D. from Leicester University in 1991 and is a Fellow of the Institution of Engineering and Technology.

Alan Dix

Computing Department

InfoLab21, South Drive

Lancaster University

Lancaster, UK

LA1 4WA

Phone: +44 (0)1524 510 319

E mail: dixa@comp.lancs.ac.uk

ALAN DIX is a Professor in the Department of Computing at Lancaster University in the UK. He was a mathematician by training, and mathematics is still his first love, but he has worked in Human-Computer Interaction since 1984, has published over 300 articles and is author of a key textbook in the area. He has worked in several universities, agricultural engineering research, local government and hi-tech start-ups. His interests are eclectic: formalisation and design, physicality and digitality, the economics of information, structure and creativity and the modelling of dreams. Recently he and a colleague have developed technology for autonomous pixels that can be configured in turn any surface or space into a two or three dimensional display.

Alex Woolley

National Centre for Product Design & Development Research, UWIC

Western Avenue

Cardiff

CF5 2YB

UK

Phone: +44 (0)29 2041 6987

E mail: a.woolley@gmail.com

ALEX WOOLLEY *is a PhD student in his final year of study. He graduated with a 1st Class Honours degree in Product Design in 2005 winning a research degree bursary later that same year. His PhD work has been concerned with the development of techniques for the rapid design and development of information appliance prototypes for testing within the context of use.*

Devina Ramduny-Ellis

Computing Department

InfoLab21, South Drive

Lancaster University

Lancaster, UK

LA1 4WA

Phone: +44 (0)1524 510 501

E mail: Devina@comp.lancs.ac.uk

DEVINA RAMDUNY-ELLIS *is a researcher in the Computing Department at Lancaster University in UK. She started her research in the area of Human Computer Interaction and Computer Supported Collaborative Work and was awarded her PhD in May 2003 for her work on developing architectural frameworks for enhancing temporal interface behaviour for distributed collaborative users. Her research interests ranges from devising methods for work analysis and interaction properties, modelling techniques and interface design, to computing infrastructure and employing empirical field studies involving ethnographic and other qualitative techniques to inform novel designs and prototypes. She has worked in several research projects within an academic environment, which covered a wide range of areas including ubiquitous computing, dependability, decision management and interfaces and infrastructure. She has over 25 publications and she is currently managing and working on the DEPtH: Designing for Physicality AHRC/EPSRC funded project. She has been actively involved in starting up a multi-disciplinary community in physicality through the Physicality 2006 and Physicality 2007 International Workshops.*

Jo Hare

National Centre for Product Design & Development Research, UWIC

Western Avenue

Cardiff

CF5 2YB

UK

Phone: +44 (0)29 2041 6987

E mail: juhare-pdr@uwic.ac.uk

JO HARE is a Product Designer and Research Assistant working on DEPtH: Designing for Physicality project in collaboration with Lancaster University and funded through the AHRC's Designing for the 21st Century call. Previous to joining DEPtH Jo worked as a product designer for a company producing medical equipment.

ABSTRACT

For some years now, the global academic and industry research communities have been working at developing techniques to rapidly design and develops information appliances such as mobile phones, MP3 players and digital cameras. Despite significant advances in the methods available for the prototyping of tangible interactive prototypes, many if not most industrial design practitioners and many UI designers still rely on two dimensional, software only interactive prototypes, particularly early in the design process when many key decisions are made. A number of attempts have been made to tackle this issue, and one of the core assumptions in many of the approaches so far taken is that designers need to be able to make "quick and dirty" prototypes in order to evaluate the tangible interactions of their concepts early in the design process. Some attempts have been made to examine how quick or how dirty the prototyping process can be for software only applications but to date no one has carried out a similar exercise for information appliance prototypes.

This paper presents the results of three separate experiments and presents empirical data that suggest answers to two important questions: Are tangible prototypes better than software prototypes?" and "how "quick and dirty" should industrial designers be aiming to prototype?"

The paper concludes by discussing the findings' significance and suggesting the implications for further work.

Keywords : *Interaction, design, fidelity levels, information appliance, product design, prototyping, design development, computer embedded products*

1 INTRODUCTION

There has been much written about the convergence of the computing, communications and media industries, particularly with reference to the advent of a particular form of tangible interactive device, the *information appliance*. Several authors have distinguished information appliances from personal computers by defining information appliances as being designed primarily to perform a specific function, whereas, in contrast, personal computers are designed to support multi-tasking[1] (Sharpe and Stenton, 2002, and Norman, 1999). Many new information appliances, such as 2.5G and 3G mobile phones, Blackberry devices, car navigation systems and new wireless music players have started appearing in recent times as a result of the convergence of the three industries.

For some years now, both academic and industry research communities have been working towards developing techniques to design and develop information appliances rapidly and efficiently, aiming to meet what Branham (2000) described as "*the need for new interactive design methods, techniques and tools to externalise thoughts and ideas, forcing the designer to be more explicit*." A number of attempts have been made to tackle this issue, among them *Wizard of Oz* simulations (Maulsby, Greenberg & Mander 1993), *Experience Prototyping* (Buchenau & Suri, 2000), *Phidgets* (Greenberg & Fitchett 2001), *Buck Method* (Pering, 2002), *Switcheroos* (Avrahami & Hudson 2002), *Augmented Reality* (Nam & Woohan 2003), *iStuff* (Ballagas et al 2003), *Paper Prototyping* (Snyder 2003), *Calder Toolkit* (Lee et al 2004), *DTools* (Hartmann et al 2006), *Exemplar* (Hartmann et al 2007) and *VoodooIO* (Villar & Gellerson 2007).

One of the core recognitions that tie these works together is that designers need to be able to make quick and "dirty" prototypes (what Schrage described as serious play (Schrage 1999) in order to evaluate the tangible interactions of their designs early in the design process. Landay and Meyers (1995) identified the value of quick and "dirty" prototyping for 2D web-based applications, their answer being *Silk* (later developed into *Denim*, Lin et al 2002), a programme that allows rapid webpage design via roughly sketched state transition diagrams linked through the exploitation of gesture recognition.

McCurdy et al (2006) made an attempt to examine *how* quick or *how* "dirty" the prototyping process can be for software only applications (using what they called mixed fidelity prototypes), but to date no one has carried out a similar exercise for tangible information appliance prototypes. Does prototyping a handheld information appliance have to involve tangible three dimensional prototyping as in the cases above? Lim et al (2006) conducted a qualitative study in this area but their investigation was focussed more on prototyping methods than fidelity levels. The tools described in earlier work such as *Toolbook* (Hustedde, 1996), *Director* (Gross, 1999) or *Hypercard* (Goodman, 1998) are all monitor-based, two dimensional systems and the derivatives of these approaches continue to be the most common methods practiced in industry. To what degree is the work developing methods for three dimensional prototypes at an early stage of product development really relevant? In other words, to what fidelity levels should industrial designers be aiming to prototype?

This paper presents empirical findings that suggest some answers. It will confine itself to examining performance, leaving more qualitative matters for future studies. Two distinct facets of physical interaction are discussed, *tangible interaction* and *physicality*. For the purposes of this paper we define tangible interaction as the interaction between a physical interface and digital information, in this case through interaction with an information appliance. *Physicality* on the other hand is a broader term which encompasses our entire interaction with the physical world. In the case of this paper we principally discuss physicality's influence through touch, feel, weight, scale etc. on our interactions with the tangible interfaces of information appliances.

2 OUR APPROACH

One of the tools the authors use is a system that allows designers to develop rapid interactive prototypes. It works by facilitating the connection of a model embedded with switches to a P.C.-based GUI prototype via a product called an *IE Unit* (Gill 2003). The system allows the P.C. to receive keyboard inputs (see *Figure 1*) so that when a user activates a switch in the model, the P.C. responds to a perceived keyboard input and a keyboard triggered GUI is activated.

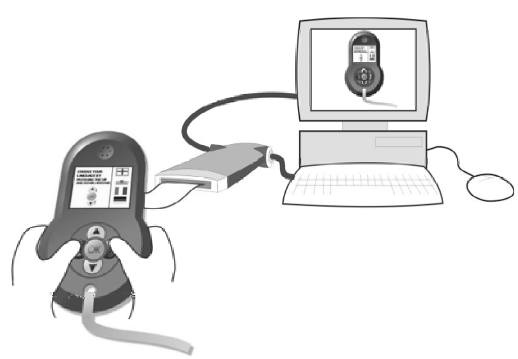


Figure 1 *Illustration shows the* IE Unit *linking a prototype to a P.C.*

[1] However some might argue that this is the result of technical limitations rather than being a design requirement.

One aspect of the system in its current form is that the display is not usually shown on the product but on a remote P.C. screen. The system is capable of facilitating models that include screens, and several prototypes have been created that include these embedded displays. However experience within the group had led to the conclusion that for the development of many types of product, this was not as important as it might appear. If this is the case, it is important as including a real screen brings with it very significant time penalties compared with an emulated screen on a P.C. Sharpe's (2002) findings were encouraging in this regard. His *Quorum* concept allowed a number of users to share digital imagery inputting in one area and receiving feedback from another.

The authors wished to find a method whereby two important questions might be answered in a quantifiable fashion:

1. Is a 3 dimensional, handheld prototype more similar to the final output than the now traditional monitor based systems most commonly used by industry?

2. How *quick* or how "*dirty*" can the prototyping process be to gain valuable feedback early in the design process, i.e. what level of fidelity is required to obtain an acceptable degree of accuracy?

The vehicle chosen for testing was the BT *Equinox* cordless phone. The authors had worked on an *IE Unit*-based prototype as part of a benchmarking exercise for a design consultancy. Part of the task of prototyping had involved mimicking the *Equinox*'s GUI interface using *Macromedia Flash*. The aim of the exercise was both to assess whether the system was capable of dealing with the complexity of a modern telecommunications interface design and to quantify to what extent it gave a true feeling of the finished interface to a potential user.

The prototype was mocked up using a set of the finished product's mouldings with its buttons wired to the *IE Unit* and a representation of the screen's output on a P.C. monitor via the *Flash* GUI.

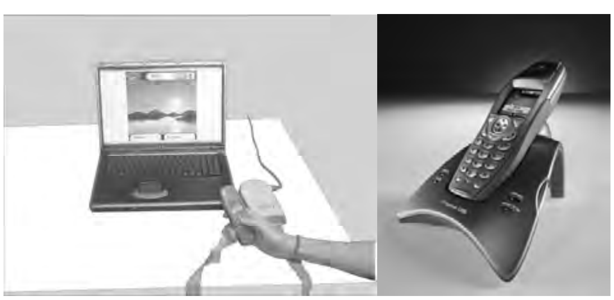


Figure 2 *On the left a mock-up of the Equinox linked to a P.C. through an* IE Unit. *On the right the* BT Equinox *phone*

In the context of the design consultancy the set-up worked effectively in that it demonstrated to the managers' satisfaction that the system was capable of producing an effective 'mock up' of a real information appliance interface by effectively mimicking the interactions of the real product. The team decided to develop the simulation further in order to carry out some empirical testing. With this in mind, some modifications were made to the mock up to enhance its functionality and a purely screen-based version of the prototype was made by modifying the way in which the *Flash* file was triggered.

HIGH FIDELITY EMPIRICAL TESTING

The team designed a programme of tests for comparing the performance of a real *Equinox* phone, the *IE Unit* prototype and the *Software* prototype. A method of conducting the tests was designed by the authors based on a methodology developed by Molich (Molich, 2002). Tasks were chosen to include common functions (ranging from simple to complex), unusual functions (such as the *Equinox*'s SMS button), and functions that involved more than straight forward transitions between the product's states.

The programme was trialled on six participants to test its effectiveness. As a result, some modifications were made to the software, hardware and methods of testing and recording data: for example, auditory feedback was added to the software simulation to confirm that a control input had been received. The team realised that this was an important aspect of the design that had to be included for a balanced trial to take place.

Experiment 1
79 undergraduate students and staff from the *University of Wales, Institute Cardiff* (UWIC) took part, ranging in age from 18 to 30 years (average age 23, 44 females and 35 males). No computer science students were included as participants, but all had at least 1 year experience using mobile phones with an average experience of 7 years. They sent an average of 6 text messages a day, suggesting good familiarity with 'typical' phone interfaces.

Procedure
Participants were divided into three independent groups (one for each manifestation of the interface, i.e. *Equinox*, *IE Unit* and *Software*) and given a series of tasks. Each participant was given an instruction sheet to read and they were allowed to ask questions if they were unsure of the procedure. They were then given one minute to familiarise themselves with the interface and technology before the tasks commenced. This was done for all participants for consistency, but was particularly important for participants using the touch screen computer (for the *Software* prototype) as this technology was unfamiliar to many. Six tasks were set for the participants. These were:

1. *Turn the phone* **on**
2. **Call** *a number*
3. **Add** *an entry to the phone's contact list*
4. *Send an* **SMS** *to a contact*
5. *Change the phone's background* **picture**
6. *Turn the phone* **off**

The six tasks were chosen because they are common mobile phone tasks. The order of the tasks was set such that the first two tasks were relatively simple so that users gained confidence using the prototype. The following three tasks were relatively complex, followed by a relatively simple task to finish. Two researchers monitored each user trial and each task was timed and graded.

The trials were also video recorded (see *Figure 3*). Comments were noted as were actions or errors of specific interest.

***Figure* 3** IE Unit *user trial*

Results

Performance of participants was converted to interval data by assigning the following numerical values to their outcome per task (0 = success, 1 = minor, 2 = serious, 3 = catastrophe). Outlying task times (3 SDs from the mean) were replaced with the next highest or lowest task times to prevent loss of data points. Two values were replaced for *On* task, 1 value was replaced for the *Call* task, 1 value for the *SMS* task and 4 values were replaced for the *Off* task. Replacements happened across all groups. Analysis of performance outcome and performance time used a 3 (device type) x 6 (phone task) mixed analysis of variance (ANOVA). The alpha level was set at .05 for significant (or reliable[2]) differences, but given the exploratory nature of these studies an alpha level of .10 was accepted as conferring marginal significance. Thus unless otherwise stated the alpha level for non-significant (or non-reliable) differences was .10. 95% confidence intervals follow reporting of means in the text.

[2] Reliable can be seen as a synonym for significant, but in common with many researchers, we tend to reserve this term for post-hoc comparisons following main or simple main effects.

Performance Time

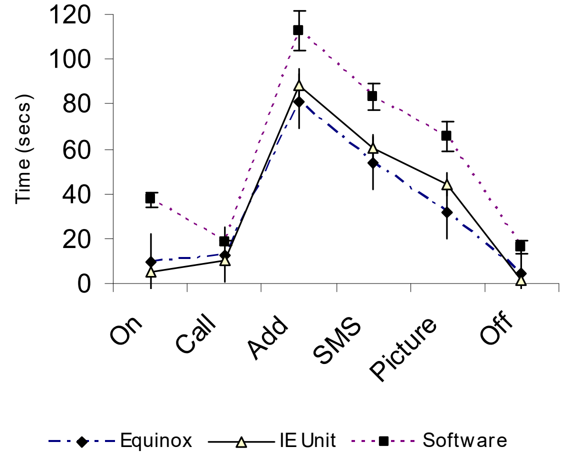


Figure 4 *Mean time taken to complete each of the six phone tasks as a function of device type. Bars are standard error.*

Figure 4 shows that the mean completion times for the software simulation were longer for all tasks, but for the other devices time difference trends were dependent on what task was undertaken. The ANOVA[3] supported this revealing a main effect of task type, $F(5, 380) = 193.08$, $p < .001$; device, $F(2, 76) = 25.01$, $p < .001$ and an interaction between task type and device, $F(10, 380) = 2.21$, $p = .04$. Post hoc tests (Tukey HSD) revealed that there were highly reliable differences ($p < .001$) between the *IE Unit* and *Software* [M = 35.08s (30.78, 39.38) vs. M = 55.54s (50.08, 61.0)] and between the *Equinox* phone [M = 32.35s (28.25, 36.45)] and *Software*. However, no reliable difference was found between the *Equinox* phone and *IE Unit*. To unpack the interaction, simple main effect analyses were undertaken looking at the difference between devices for each type of task. These showed that for every task there was a significant difference between devices (smallest F = 4.18, largest p = .019). Subsequent simple comparisons (Tukey HSD corrected) showed that with the exception of *Add to phonebook* task there were reliable pairwise differences ($p < .05$) between *IE Unit* and *Software*, *Equinox* and *Software* but no reliable pairwise differences between IE Unit and *Equinox* for any of the tasks. In the case of the *Add* task, *Equinox* was reliably different from *Software* ($p < .05$) but *IE Unit* was marginally different from software ($p = .06$). As per the other tasks *Equinox* and *IE Unit* did not differ reliably.

[3] For this and subsequent analyses the main points of interest are a) whether significant differences exist between the prototyping devices and b) whether these differences interact with the type of task undertaken. Due to the 'unique' nature of each task significant task differences by themselves are very much a secondary concern and hence not subjected to follow up analysis.

Thus the interaction suggests that whilst *IE Unit* and *Equinox* are more alike than the *Software* solution (in terms of time taken) on each task, the magnitude of this effect is mediated by the type of task undertaken.

Performance Rating

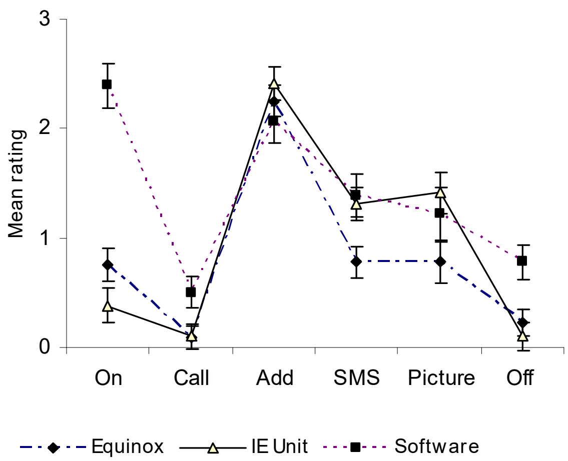


Figure 5 *Success outcome (rating) for each of the six phone tasks as a function of device type. Bars are standard error.*

Figure 5 shows that *Equinox* and *IE Unit* were more similar for most tasks than the *Software*, but that over all three devices task specific differences did exist. The ANOVA supported this revealing a main effect of task type, F (5, 380) = 62.43, p < .001; device, F (2, 76) = 12.34, p < .001 and an interaction between task type and device, F (10, 380) = 7.47, p < 001. Post hoc tests revealed that there was a highly reliable difference (p < .01) between *IE Unit* and *Software* [M = .95 (.81,1.10) vs. M = 1.39 (1.20,1.58)] and between the *Equinox* [M = .81 (.67, .95), p < .001] and *Software*. However, no reliable difference was found between *Equinox* phone and *IE Unit.* Simple main effects analyses showed that there were significant differences between devices for the *On*, *SMS* and *Off* tasks (smallest F = 4.40, largest p = .02), and marginally significant differences between the devices for the *Call* and *Picture* tasks (smallest F = 2.89, largest p = .06) but no significant difference existed between the devices for the *Add* task. Subsequent simple comparisons revealed that for the *Off* and *On* tasks there was a reliable (p < .05) difference both between *IE Unit* and *Software* and between *Equinox* and *Software*. For the *Call* task there were marginally reliable differences between *Equinox* and software (p = .07) and between *IE Unit* and *Software* (p = .08). For the *SMS* task there were reliable differences (p < .05) between *Equinox* and *Software* and *Equinox* and *IE Unit*. For the *Picture task* there was a marginally reliable (p = .06) difference between *Equinox* and *IE Unit*. None of the other task by device comparisons showed reliable differences. Thus the interaction shows that whilst *IE Unit* and *Equinox* exhibit similar success patterns on some tasks there are others where they can be quite dissimilar. It is important to note however that in the two areas where significant differences occurred, the *IE Unit* was similar to *Software*. In other words, the tangible model is never significantly worse than a software only simulation, but in some tasks offers far closer results to the real device.

The above analyses show that, on both the time taken to complete a task and on how successfully it performed, the *IE Unit* tangible prototype was more like the real *Equinox* phone than the *Software* simulation. Nevertheless, there were also significant interactions between the different phone tasks and the performance measures, particularly those measuring 'success' at completing the tasks.

The exceptions are the *Call* and *Picture* tasks. Although we do not have hard data, we can speculate as to why these may have been different.

The tasks fall into three main types:

(i) The *On* and *Off* tasks which require finding a physical button on the phone, but do not require viewing of the screen, except maybe to confirm it has turned on/off.
(ii) The *Call* task which is mainly concerned with typing a number, then possibly checking the number on the screen before locating the 'call' button. Like (i) this is predominantly an 'eyes down' task looking at the device, except that the keys to press are more obvious.
(iii) The *Add*, *SMS* and *Picture* tasks which all require divided attention between the device (eyes down) and the screen (eyes up).

The problems with (i) are discussed below based on participant comments. The *Call* task (type ii) involves either hitting buttons, or pressing clearly identifiable buttons on screen, both of which are straightforward actions and specifically do not involve any attention switching. Both *SMS* and *Picture* tasks are of type (iii) where the user has to switch attention between device and screen during interaction.

There is a strand of research looking at the way personal devices such as phones or PDAs can be used to interact with larger displays (Sas and Dix, 2008). In one of these studies, Gostner et al. (2008) found that users indeed appear to perform acceptably with the divided attention and yet still comment on the problems it causes them. The higher rate of outcome problems with these tasks when the *IE Unit* is used may be due to this attention switching. This is not evident in the *Add* tasks, but it is reasonable that errors or difficulty due to attention switching are related to fine details of the task.

Discussion

The data makes it clear that the *Equinox* and the *IE Unit* performed in a more similar fashion than the software alone. This is significant because the *IE Unit* is intended as a design tool to prototype and test tangible user interfaces.

The more variance from the results of the *IE Unit* prototype versus the actual product (whether the *IE Unit* results are more successful or less), the less effective a tool it will be. As has been demonstrated, the *IE Unit* produced a consistently more realistic simulation than software alone, which opposes the claims made by Sharp (1998). In Norman's (1988) theorizing, the system image created by the *IE Unit* is a better fit of the user's mental model of a phone device than a purely software simulation. This result is all the more significant for two major factors:

a) A phone is a ubiquitous information appliance (Weiser, 1994) and all participants had experience with similar devices.

b) The chosen appliance had a push button interface with all its controls mounted on the top surface. This combination allows the software prototype to compete on favourable terms with the other methods. It is of course conjecture at present, but had the selected appliance featured sliders, dials, triggers etc, or had the controls been mounted in a more three dimensional fashion around the product, then the software simulation may have matched the performance of the real product even less.

As mentioned, it is likely that many learnt phone tasks place more demands on implicit than explicit memory systems (cf. Graf & Schacter, 1985). When solving novel tasks people will draw upon previous experience or schemas (Anderson, 2005). We might hypothesise then that participant's *phone schemas* contained much information in a motoric (and implicit) representational format. Thus when completing the phone tasks used in the current studies, the *IE Unit* afforded better use of past experience, as the need for physical interaction effectively served as a memory trigger for this schematic knowledge.

Nevertheless, designers do need to exercise some caution. Good as the tangible prototype's performance was overall, it did simulate some tasks better than others. Thus, further work now needs to be carried out to ascertain how very rapidly conceived and prototyped three dimensional appliances designed using the latest techniques perform against those prototyped using traditional methods.

Participants were encouraged to comment on their experiences. Some of these comments have been included below to illustrate certain assertions.

Software simulation:
There were two highly visible issues with the *Software* method. The first was with the *Power On* and *Power Off* tasks. Participants repeatedly struggled to detect the location of the switch. A very common error was to press the power symbol as opposed to the switch which was situated to the left of the symbol (see *Figure 6*): "*The only problem I had was switching it on…the power button is much clearer on the real phone*".

Figure 6 *Screen shot from software* Equinox *simulation showing power button highlighted*

The second conspicuous issue was that participants struggled to differentiate buttons from areas of the screen: "*I kept on pressing the words on the screen but I suppose you wouldn't if you had the actual phone*". "*Because the screen icons looked like buttons I tried to press them*". "*I keep wanting to press the symbols on the screen. It's 'cause I don't have the phone in my hands*". One participant stated: "*Technology gets in the way*".

These two problems were not observed in the tangible prototype because physical buttons have perceived affordances (Norman, 1988), that is they expose aspects of their potential behaviour through their physical appearance – screens, buttons and labels are all different. Furthermore if one does mistakenly press a non-button it is immediately obvious as it does not depress. In contrast an on-screen button and a labelling icon can look very similar, and the error of pressing the latter is only apparent in the semantic feedback of the phone not going on.

While in the experiment this was a problem in the fidelity of emulation, it is not uncommon to see flat membrane buttons on physical devices leading to similar problems. Indeed, one of the authors was once trapped in a train toilet until he realised that the label saying 'press to unlock' was not a label, but in fact the button!

It appears then that a participant's mental model of phone interaction (derived from using a physical device) was not fully transferable to the touch screen software simulation. Thus, learning task-action mappings in one interaction domain does not necessarily transfer readily to another.

***IE Unit* simulation:**
On the whole users commented on how closely the tangible *IE Unit* prototype simulated the real device: "*Similar to the real product to use*". "*Quite straight forward. Simulation fairly good, no problems*". One user commented that they: "*Find simulation quite easy, had some problems with the raised buttons*". This was an important comment because while it is in one sense negative, in the 'field' the outcome would be positive, i.e. designers would have discovered a potential issue with the button design. The software method is not able to do this. As theorists such as Norman (1988)

have highlighted, the feedback provided from buttons, dials, sliders etc is a crucial determinant in product usability.

There were two issues raised with the *IE Unit* method however. One user complained that it was: "*More difficult using IE simulation – affected entry of text a little bit*"[4]. A clearer issue was the fact that the interaction with the appliance and screen are separated. One participant summed this up: "*More convenient if this was the real phone – had to look from the phone to the (P.C.) screen – this affected text messaging*". This aspect is significant. The authors' research has found that some in industry view the screen being included as a strong necessity (particularly mobile phone designers) while others prefer the simplicity and flexibility of maintaining a discrete screen.

LOW FIDELITY EMPIRICAL TESTING

So far we have seen that at high fidelity levels, a hand held product linked to a computer simulation gives data that is of higher quality than the standard screen-based industry method of simulation. This is certainly useful but the benefits must be balanced against the extra work and therefore cost of creating a high fidelity prototype connected to the simulation. Building a prototype at high fidelity might double the time and cost of creating the simulation and one must therefore ask whether the benefits of a more accurate representation of the end results are enough to outweigh the time and fiscal penalties.

The question now arose however as to how much effect the physical interaction was having on the user. In other words, could a lower fidelity model and interface give useful results in the same or less time than an entirely screen-based prototype? The authors decided to run more empirical tests in order to test the hypothesis that physical interaction with a prototype was more important than the fidelity level of either the model or the interface.

Low Fidelity Modelling

The team elected to continue using the *Equinox* for further tests so that direct comparisons could be made between the data gathered in the high and low fidelity testing phases. A new model was accordingly produced from "soft" modelling materials. The main body of the phone was constructed in blue foam (a standard product designer's soft modelling material) with the switches being topped with card cut-outs in the shape of the switches on the real phone. On top of these were glued the button graphics, and the screen was represented by a piece of coloured paper.

The modelling process took around one working day to complete including embedding the switches. A further working day was expended creating a new, low fidelity GUI in *Flash*. The new *Flash* GUI was created using sketch work produced on screen through the mouse. The GUI was driven via keystrokes in the same way as the higher fidelity prototype described above.

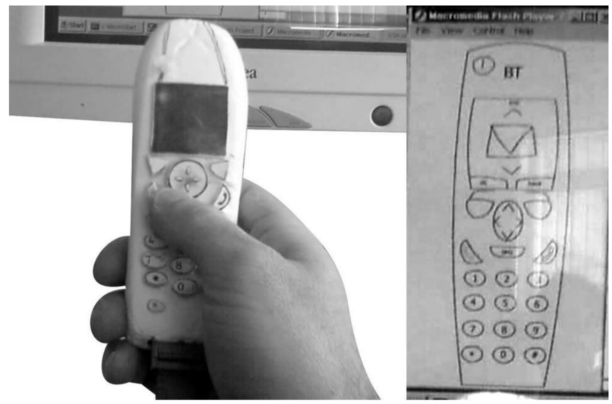


***Figure* 7** *Low fidelity Equinox model connected to low fidelity GUI (screen shot shown on right)*

The low fidelity prototype's hardware was identical to that of the higher fidelity model, indeed it worked as well as the high fidelity GUI described in the first section though it was produced with reduced functionality which decreased the mock up time of the interface. Nevertheless, the reduced hardware production time meant that for equivalent levels of functionality the low-fidelity GUI would still give a time-saving of around 4 days compared to the higher fidelity GUI. The significance of these figures is that the low fidelity *Equinox* prototype linked to the low fidelity GUI was manufactured in 20% of the time it would take to create the high fidelity touch screen interface with equivalent functionality. If, therefore, this method was found to produce results similar to the real product, then the viability of rapidly designed and produced three dimensional prototypes would have been proved.

The team therefore set out to test the effectiveness of the low fidelity setup using exactly the same testing methods as before but with a more limited set of tasks for speed.

Experiment 2

16 undergraduate students and administrative staff from the *University of Wales, Institute Cardiff* (UWIC) took part ranging in age from 18 to 30 years. Experience of mobile phone interfaces was broadly similar to that in Experiment 1[5].

[4] It should be noted that this user failed in the first text entry task, although his time in the second was half the average for others using the *IE Unit*. It is therefore difficult to draw any strong conclusions.

[5] Although slightly repetitive, complete ANOVA analyses are given despite the fact that many of the trends reported previously will obviously still remain, e.g. IE Unit versus Software versus Equinox comparisons. However, the addition of additional prototypes does increase the degrees of freedom and thus may have altered the pattern of previous findings.

The procedure was similar to Experiment 1 except that data was not collected for the *SMS* and *Add* task. Apart from this all other conditions were identical allowing comparison between this new data on the new 'low fidelity' interface and the data from Experiment 1. In the discussion below, the 'low fidelity' interface is referred to as *Sketch*.

Performance Time

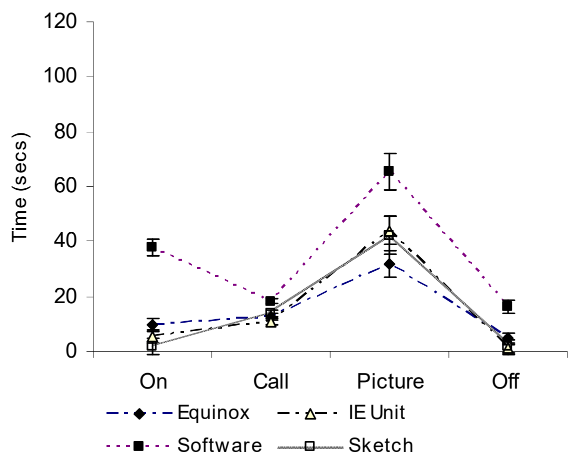


Figure 8 *Mean time taken to complete each of the four phone tasks as a function of device type. Bars are standard error.*

Figure 8 shows that the mean completion times for all tasks except *Call* were slower with *Software* than the other prototypes. The ANOVA supported this revealing a main effect of task type, F (3, 273) = 123.77, p < .001; a main effect of device F (3, 91) = 19.06, p < .001; and an interaction between task type and device, F (9, 273) = 4.69, p < .001. Post hoc tests revealed that *Software* was highly reliably slower (p < .001) than all the other devices [M = 34.38s (29.79, 38.96) versus M = 14.73s (11.29, 18.17), 15.34s (11.72, 18.95), 14.88s (10.0, 19.74) for *Equinox*, *IE Unit*, *Software*, and *Sketch* respectively] but that the other devices were not reliably different from each other. Simple main effects analyses showed that there was a significant difference between the devices for each of the 4 tasks (largest F = 32.53, largest p = .001). Simple comparisons showed that for *On* and *Off* tasks *Software* was highly reliably slower (all ps < .005) in response times than each of the other devices, although these other devices were not reliably from each other. For the *Call* task *Software* was reliably slower (all ps < .05) than *Equinox, IE Unit* and *Sketch.* For the *Picture* task *Software* was reliably different (p < .05) than *Equinox* and marginally reliably different (p < .08) than both *IE Unit* and *Sketch*. For all of the devices and tasks, none of the other simple comparisons were reliably different.

Thus, in terms of the time taken to complete the tasks, it seems that in general, the software simulation was slower, whilst the physical prototypes were more similar to each other and to the actual phone. However, this pattern was not of the same magnitude for all tasks, so this needs to be considered when interpreting the findings, i.e. physical prototypes (versus software) appear to simulate most but not all tasks better.

Performance Rating

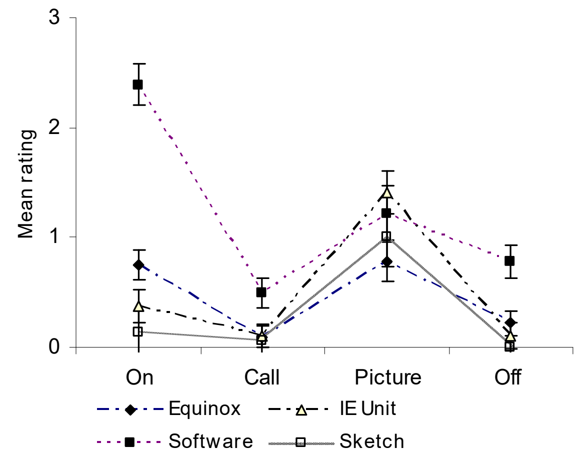


Figure 9 *Success outcome (rating) with four phone tasks as a function of device type. Bars are standard error.*

Figure 9 suggests that overall performance of the low fidelity prototype *Sketch* was more similar to the *Equinox* than the software simulation, but that this trend was more apparent for some tasks and less for others. The ANOVA supported this revealing a main effect of task type, F (3, 273) = 34.29, p < .001; device F (3, 91) = 15.19, p < .001 and an interaction between task type and device; F (9, 273) = 7.60, p < .001. Post hoc tests showed that the *Equinox* [M = .46 (.30,.62)], *IE Unit* [M = .50 (.33, .67)] and *Sketch* [M = .30 (.07, .52)] were all highly reliably different (p < .001) than *Software* (M = 1.22 (1.01, 1.43)] but none was reliably different from each other. Thus on the 4 tasks chosen the low fidelity *Sketch* prototype was more similar in success rating to the real product and the higher fidelity 'physical' prototype than it was to the software simulation. Simple main effects analyses showed that there were highly significant differences between the devices for both the *On* and *Off* tasks (largest F = 30.83, largest p = .001) and a marginally significant difference (p = .06) between the devices for the *Call* task. However the devices did not differ significantly from each other for the *Picture* task. Simple comparisons showed that for the *On* task, *Sketch*, *IE Unit* and *Equinox* were highly reliably different (p < .001) than *Software* and that *Sketch* was marginally different (p < .06) than *Equinox*, whilst for the *Off* task all devices were reliably different (p < .05) than *Software* but not reliably different from each other. For the *Call* task, *Equinox* and *IE Unit* were marginally reliably different from each other (all ps < .10)

Thus it seems that the behaviour of the low fidelity *Sketch* prototype is most similar to that of the higher fidelity *IE Unit* prototype and the *Equinox* device itself. In the case of success outcome for the *Picture* task there is a less clear effect. However, if we look at those conditions where the software prototype differed substantially from the real *Equinox* device, specifically the *On* and *Off* tasks and even the performance data for the *Picture* tasks, it is evident that in these tasks *Sketch* is very similar.

Discussion

The software simulation continues to perform badly in comparison to the *Sketch* model. Generally speaking *Sketch* continues to demonstrate the importance of physicality in gaining accurate results. Curiously it actually very marginally outperforms the higher fidelity *IE Unit* prototype. In any case the significance of these results lie in the fact that more accurate results were produced from a “quicker, dirtier” tangible prototype produced with an 80% time saving over a high fidelity screen-based interface.

***Figure* 10** Flat Face *low fidelity* Equinox *model: blue foam with embedded switches covered by a printed sheet of paper*

Further Fidelity Reduction

After reviewing this data the authors decided that it would be useful to investigate whether lowering the fidelity level further would maintain the tangible prototype’s performance edge over the virtual.

A still lower fidelity tangible prototype was constructed. Like the other low fidelity unit the main body was constructed from blue foam. This time however, instead of modelling the front face, a full size print out of a front view of the Equinox phone was glued over the tops of the buttons. (see *Figure 10*) The paper allowed enough flex so that when the user pressed on a picture of a button the paper the real button under it was activated. There are three important factors that should be noted about this approach:

1. The user should, in theory have no more clues as to functionality of this kind of tangible prototype than with the wholly screen based prototype.

2. Notwithstanding that fact, when a control is activated the user does receive tactile feedback in a way that the screen based prototype does not allow.

3. Other physical interactions are similar to the real phone and the other tangible prototypes.

Experiment 3

16 undergraduate students and administrative staff from the *University of Wales, Institute Cardiff* (UWIC) took part ranging in age from 18 to 30 years. Experience of mobile phone interfaces was broadly similar to that in *Experiment 2* and the procedure was identical, again allowing comparison with data from *Experiments 1* and *2*. The same low fidelity GUI was used as in *Experiment 2* with only the prototype itself changed as described above. This prototype is referred to as *Flat face* in the following analysis and discussion.

Performance Time

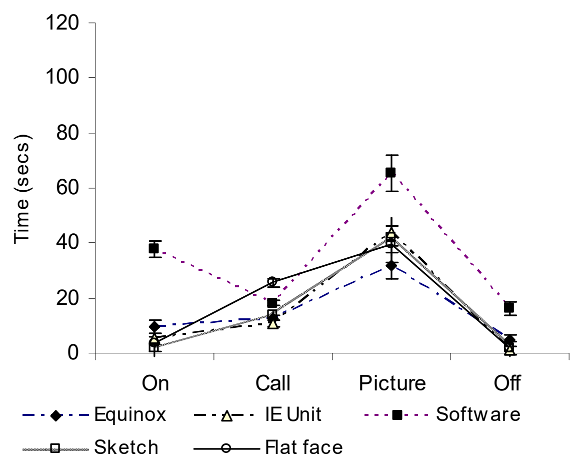


Figure 11 *Mean time taken to complete each of the four phone tasks as a function of device type. Bars are standard error.*

Figure 11 shows that the mean completion times for all tasks except *Call* (*Flat face* was slowest here) were slower with *Software* than the other prototypes. The ANOVA supported this revealing a main effect of task type, $F(3, 318) = 148.86$, $p < .001$; a main effect of device $F(4, 106) = 15.58$, $p < .001$; and an interaction between task type and device, $F(12, 318) = 5.17$, $p < .001$. Post hoc tests revealed that *Software* was highly reliably slower ($p < .001$) than all the other devices but that the other devices were not reliably different from each other. Simple main effects analyses showed that there was a significant difference between the devices for each of the 4 tasks (largest $F = 30.28$, largest $p = .002$). Simple comparisons showed that for *On* and *Off* tasks *Software* was highly reliably slower in response time than each of the other devices, although these other devices were not reliably different from each other. For the *Call* task *Software* was reliably slower than the *Equinox, IE Unit* and *Sketch,* but reliably faster than for *Flat Face*. *Flat Face* was reliably slower (all ps < .02) than every other device. For the *Picture* task, *Software* was reliably different (p < .05) than *Equinox* and *Flat Face* and marginally reliably different (p < .09) than *IE Unit* and *Sketch*. For all of the devices and tasks none of the other simple comparisons were reliably different.

These results are very similar to those of *Experiment 2*. Whilst the *Flat face* prototype is, from the front, visually identical to the on-screen software interface, its behaviour is still very similar to the higher-fidelity prototypes. The notable difference is the slower time for the *Call* task. Given this is the most important function on the phone, this is not an unimportant difference! This reminds us that all results from prototypes need to be regarded with an element of caution. In addition, the *Flat face* prototype is in some ways similar to touchscreen-based phones and consumer devices such as the *iPhone*, suggesting that care needs to taken in designing such devices in order to ensure they are usable as well as desirable.

Performance Rating

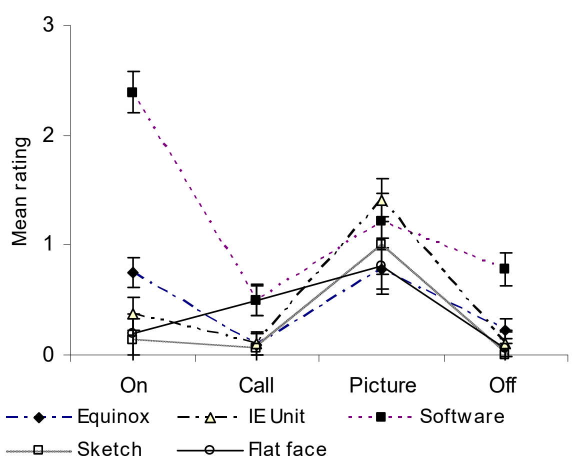


Figure 12 *Success outcome (rating) with four phone tasks as a function of device type. Bars are standard error.*

If we consider the outcome rating (*Figure 12*) we see a similar picture to the data for Performance Time. The ANOVA reveals a main effect of task type, F (3, 318) = 31.82, p < .001; device F (4, 106) = 12.76, p < .001 and an interaction between task type and device. Post hoc tests showed that *Equinox*, *IE Unit*, *Sketch* and *Flat face* were all highly reliably different (p < .001) than *Software* but neither was reliably different from each other. Again, simple effects analyses were used to unpack the interaction. These showed that there were highly significant differences between the devices for both the *On* and *Off* tasks (largest F = 28.45, largest p = .001) and a significant difference (p < .05) between the devices for the *Call* task. However the devices did not differ significantly from each other for the *Picture* task. Simple comparisons showed that for the *On* task, both *Sketch* and *Flat face* were highly reliably different (p < .001) than the *Software* simulation and that *Sketch* was marginally reliably different (p < .06) than *Equinox*, whilst for the *Off* task all devices were reliably different (p < .05) than *Software* but not reliably different from each other. For the *Call* task none of the simple comparisons were significant, suggesting that none of the devices were reliably different from each. This is despite the patterns that appear evident in *Figure 12*. However, when LSD (unadjusted)[6] simple comparisons were used, *Flat Face* and *Software* were again not reliably different from each other, though *Flat Face* was reliably worse than all the other devices (all ps < .05) and *Software* was also reliably worse than all the other devices (all ps < .05).

As with the task time measures, *Flat face* has very similar behaviour to the real device and higher-fidelity prototypes, with the exception of the *Call* function, where, like the *Software* interface, we see more errors/problems.

It was surprising that Flat face did not have similar problems in tasks *On* and *Off* to those of the *Software* prototype. Recall that users tried to click the label on screen rather than the actual power button. However, in re-examining the *Flat face* prototype, the reason for this became clear. Because the paper covering has a certain stiffness, if the power icon is pressed the button next to it is in fact activated. This certainly emphasises that small differences in physical materials can make a significant difference in behaviour. This is important when choosing physical materials to use in prototyping, as is the case here, but also in selections for the product itself – in this case a form of membrane keypad has changed the effective 'size' of buttons.

Discussion

The results of *Experiment 3* were surprising in some instances. It was noted in *Experiment 1* that the *Power On* and *Power Off* tasks were problematic on the purely screen based prototype because users found it difficult to find the switch on the screen. In *Experiment 3* the front of the tangible prototype was formed by a piece of paper on which was printed the exact image used in the screen based prototype. So the visual clues available to the user were the same. The authors at first concluded that physicality must be playing a subtle role in the flat faced tangible prototype for this function. This may in fact be true, but as noted above, the user could in fact make an error with the *Flat face* prototype that on *Software* would have resulted in failure. For the *Call* function the *Flat face* prototype performed

[6] Some researchers routinely use unadjusted simple comparisons following a significant simple main effect (as is the case with *Call*). However, as the other analyses in this paper have chosen to use a harsher criterion, the results must be treated more cautiously.

similarly to the *Software* prototype and significantly worse than the *Equinox, IE Unit* and *Sketch* prototypes. The authors concluded that in this task where several numbers on the keypad needed to be pressed (rather than just the on/off button or navigation button in other tasks) the flat face of the prototype did not replicate the true physicality of the product sufficiently and the result was more user error resulting in slower performance times and worse performance ratings. It may also be that when a user chose a space between the actual buttons the paper tension was causing adjacent controls to activate. User comments during the tests with fully 50% of participants in the *Flat face* trials reveal that their frustration with the *Flat face* prototype, making specific complaints about the quality of the prototype *at the point where they had to activate a number of controls in a sequence* (dialling a number).

	material	**screen**	**keypad**	**behaviour**
Equinox	production mouldings	on device	raised keys	–
IE Unit	production mouldings	on P.C.	raised keys	mostly similar to *Equinox*, except success rates lower on split attention tasks
Sketch	blue foam	on P.C.	paper capped	similar to *IE Unit*
Flat face	blue foam	on P.C.	flat paper	similar to *IE Unit* & *Sketch* except *Call* success rate more like *Software*
Software	virtual	on P.C.	virtual	slower on all tasks except *Call*, and lower success rate for all tasks

Table 1 *Summary of Prototypes used in Experiments*

CONCLUSIONS

The authors started this paper by raising the question of the speed of prototyping and prototype fidelity levels industrial designers should be aiming for. The results of the experiments above would suggest it is not the level of fidelity that is most important but rather considerations of tangibility and physicality. The extent to which tangible prototypes of hand held information appliances appear to outperform screen based prototypes in the simulation of an actual product were perhaps its most unexpected features. The fact that the advantage continues even when the tangible prototypes are made five times as quickly and at much lower fidelity levels underscores the issue.

The findings' significance therefore lie in the fact that there would appear to be merit in the adoption of tangible prototyping methods, particularly at low fidelity levels. However, if these low fidelity tangible prototypes compromise on the physical attributes of the design, such as removing the tactile feedback of buttons (as was the case for the *Flat face* prototype), this could affect user performance, significantly. The degree to which performance is affected alters dramatically according to task type (*Table 1* summarises the main differences in behaviour). Physicality clearly plays an important role in users' interaction with handheld products but the authors were surprised at the extent to which even very subtle tactile feedback such as the switches under the taut surface of a piece of paper appeared to make a very marked difference to users' ability to interact smoothly with the product. Clearly in some cases it leads to apparently positive results (good response times for *Power On* or *Off*) and sometimes poor performance (the number of users who complained that they were unable to work the prototype because it wasn't obvious enough where the buttons were). In any case, the authors are of the view that the *Flat face* prototyping technique brought with it more compromises than the method used for *Sketch*.

All of these factors create challenges for those researchers developing toolkits for the development of tangible information appliance prototypes. What is really needed by the design community is a toolkit allowing the flexibility and speed of *Paper Prototyping* (Snyder 2003) or *Wizard of Oz* (Maulsby et al 1993), the software integration of *DTools* (Hartmann et al 2006), the wireless capability and flexibility of input trigger placement of the *Calder Toolkit* (Lee et al 2004) and the ability to use off the shelf components of the *IE System* (to exploit a wide range of appropriately scaled input mechanisms) as described in Gill (2003).

Hartmann et al (2006) noted the importance of a small form factor and this work emphasises that aspect. In essence, if the physicality of tangible prototypes is important as this work suggests it might be, then it follows that scale is likely to be an equally important issue since key aspects of any interaction with a handheld device are heavily dependent on size and control input groupings. To date, most technology based toolkits have only been capable of prototyping oversized representations of the appliances they represent. Further work is needed to determine whether the scale of a tangible prototype has a significant bearing on the accuracy of its simulation of a real information appliance.

The purpose of the experiments described in this paper was to understand the role of tangible prototypes and physical fidelity in the design process. However, as a side effect it has also studied what could be regarded as a range of separate interfaces to the same underlying functionality but with a range of different physical forms. Tangible interfaces and devices are often compared with a 'normal' interface with equivalent functionality, but the differences are typically large and cover many factors making it hard to trace precise causes. In contrast we have a number of quite fine physical distinctions and can observe where these either have no effect, or where there was an effect, precisely which change caused it. Thus at various stages we have highlighted potential wider implications.

As an example of this, in *Experiment 1* we saw that even problems with split attention tasks were very dependent on the precise balance of the task. This has important implications for those experimenting with systems for using mobile phones to interact with public displays (Sas and Dix 2008), or even designing television remote controls. Experimental tasks need to be carefully designed in order to cover, not only a range of eyes down and eyes up tasks, but also variants of each.

This study has dwelt primarily on quantitative methods, and for reasons of brevity and focus has not explored more qualitative aspects in depth. Future studies may benefit from a more qualitative approach. One of the aims of the AHRC/EPSRC funded DEPtH project, of which most of the authors are members, is to explore these issues in detail, arriving at conclusions regarding the level of fidelity for which designers should be aiming in prototype work. This exploration includes a variety of methods including qualitative ethnographic and content analysis techniques, and formal modelling of physical devices, as well as quantitative experiments.

ACKNOWLEDGEMENTS

This work has been supported by the AHRC/EPSRC funded project DEPtH “Designing for Physicality” (http://www.physicality.org/), part of the joint AHRC/EPSRC research “Designing for the 21st Century” Initiative (http://www.design21.dundee.ac.uk/).

REFERENCES

Anderson, J.R. (2005). Cognitive Psychology and its implications (6th ed). *New York, Worth.*

Avrahami, D., Hudson, S. (2002) Forming Interactivity: A Tool for Rapid Prototyping of Physical Interactive Products", *in Proceedings of DIS'02, pp. 141-146*

Ballagas, R., Ringel, M., Stone, M. and Borchers, J. (2003) iStuff: A physical user interface toolkit for ubiquitous computing environments *in Proceedings of CHI 2003, pp. 537-544*

Branham, R. (2000). Given the radically changing work environment and new worldviews, what kinds of new ‘tools’ do designers need to survive and successfully deal with tomorrow’s design problems? *Proceedings of IDSA 2000 National Education Conference*, Louisiana, USA.

Buchenau, M. and Suri, J.F. (2000) Experience Prototyping. In *Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques*, 424 – 433, New York City, New York, USA.

Gill, S. (2003) Developing Information Appliance Design Tools for Designers. In *Proceedings of the 1st Appliance Design Conference*, Bristol, UK.

Goodman, D. (1998), *Complete HyperCard 2.2 Handbook*, Lincoln: iUniverse.

Gostner, R. Gellersen, H., Kray, C., Sas, C. (2008). Reading with Mobile Phone & Large Display. in *Proceedings of Workshop on Designing and Evaluating Mobile Phone-Based Interaction with Public Displays, CHI 2008*, Florence, April 2008.

Graf, P., and Schacter, D.L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 11*, 501-518.

Greenberg, S., Fitchett, C., (2001) Phidgets: easy development of physical interfaces through physical widgets, in *Proceedings of the 14th annual ACM symposium on User interface software and technology*, Orlando, Florida, USA 209 - 218

Gross, P., Tucker, K. (1999) *Macromedia DIRECTOR 7 and LINGO AUTHORIZED*, San Francisco: Macromedia Press.

Hartmann, B., Abdulla, L., Klemmer, S., and Mittal, M. (2007) Authoring Sensor-based Interactions by Demonstration with Direct Manipulation and Pattern Recognition in *Proceedings of ACM CHI 2007 Conference on Human Factors in Computing Systems*

Hartmann, B., Klemmer, S.R., Bernstein, M., Abdulla, L., Burr, B., Robinson-Mosher, A., Gee, J. (2006) Reflective physical prototyping through integrated design, test, and analysis in *Proceedings of UIST 2006*

Hustedde, S.F. (1996), *Developing With Asymetrix Toolbook : Applied Programming Theory*. Belmont: Wadsworth.

Landay, J.A. and Myers, B.A. (1995) Interactive Sketching for the Early Stages of User Interface Design *in Proceedings of CHI '95*

Lee, J. C., Avrahami, D., Hudson, S.E., Forlizzi, J., Dietz, P.H., & Leigh, D. (2004) The Calder Toolkit: Wired and Wireless Components for Rapidly Prototyping Interactive Devices, in *Proceedings of Designing Interactive Systems Conference*, Cambridge, Massachusetts.

Lim, Y, Pangam, A., Periyasami, S. and Aneja, S. (2006), *Comparative Analysis of High- and Low-Fidelity Prototypes for More Valid Usability Evaluations of Mobile Devices* in Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles, Oslo, Norway, 291 – 300

Lin, J., Thomsen, M. and Landay, J.A. (2002) A visual language for sketching large and complex interactive designs in *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*

Maulsby, D., Greenberg, S. and Mander, R. (1993), Prototyping an intelligent agent through Wizard of Oz, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, Amsterdam, The Netherlands

McCurdy, M., Connors, C., Pyrzak G., Kanefsky B., and Vera, A. (2006) Breaking the Fidelity Barrier: An Examination of our Current Characterization of Prototypes and an Example of a Mixed-Fidelity Success in *Proceedings of CHI 2006 Usability Methods, Montréal, Québec, Canada*

Molich, R., Dumas, J. S. (2006) Comparative usability evaluation (CUE-4) *Behaviour & Information Technology, 1 – 19, December 2006, Taylor & Francis*

Nam, T.J. & Woohan, L. (2003) Integrating Hardware and Software: Augmented Reality Based Prototyping Method for Digital Products in *Proceedings of CHI 2003 Conference on Human Factors in Computing Systems Fort Lauderdale, Florida.*

Norman, D.A (1988). *The Psychology of Everyday things*. New York: Basic Books.

Norman, D.A. (1999). *The Invisible Computer*, MIT Press.

Pering, C. (2002) Interaction Design Prototyping of Communicator Devices: Towards meeting the hardware – software challenge. *Interactions journal* 9(6) 36-46.

Sas, C., Dix, A. (2008). Designing and evaluating mobile phone-based interaction with public displays. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 3941-3944.

Schrage, M. (1999) *Serious Play: How the World's Best Companies Simulate to Innovate*, Harvard Business School Press.

Sharp, J. (1998) *Interaction Design for electronic products using virtual simulations*, Unpublished doctoral dissertation, Brunel University, UK.

Sharpe, W.P. and Stenton, S.P. (2002) Information Appliances. In J. Jacko, A. Sears (Eds.) *The Human Computer Interaction Handbook*, (pp.731 – 751). Mahwah: Lawrence Erlbaum Associates

Snyder, C. (2003), Paper Prototyping, Morgan Kaufmann, San Francisco, CA.

Villar, N. and Gellerson, H. (2007) A Malleable Control Structure for Softwired User Interfaces in *Proceedings of Tangible and Embedded Interaction Conference (TEI '07)*

Weiser, M. (1994). The world is not a desktop. *Interactions,* ACM, 1 (1), 7-8.

Physical Fidelity: Exploring the Importance of Physicality on Physical-Digital Conceptual Prototyping

Joanna Hare[1], Steve Gill[1], Gareth Loudon[1], Devina Ramduny-Ellis[2], and Alan Dix[2]

[1] The National Centre for Product Design and Development Research (PDR), The University of Wales Institutue, Cardiff (UWIC), Western Avenue, Cardiff, CF5 2YB, UK
{juhare-pdr,sjgill,gloudon}@uwic.ac.uk
[2] InfoLab21, Lancaster University, Lancaster, LA1 4WA, UK
devina@comp.lancs.ac.uk, alan@hcibook.com

Abstract. The physicality of digital-physical devices is an essential part of our interaction and understanding of information appliances. This paper draws on the findings of an empirical study investigating the effect of physical fidelity on a series of user trials. Three prototypes of a single design intent were built, the standard of their construction dictated by the time imposed on the designer. In choosing this constraint, the authors present the argument that the most important driver in decisions that dictate fidelity levels is the available and/or necessary time required for making a prototype in order to generate information of the right quality. This paper presents the empirical and qualitative results of the trials, which suggest that there is little effect of fidelity on user performance, but the user's ability to give constructive feedback on the design was influenced by the nature of the prototypes.

Keywords: Physicality, prototyping, fidelity, information appliance, product design, tangible interface, low fidelity prototyping.

1 Introduction

This paper focuses on *information appliances,* devices designed to do one task, but do it well. The design of these devices poses interesting challenges to the design community because not only do information appliances have physical considerations (size, shape, buttons, etc.), they also have digital considerations (dedicated computer running software menus, features, function etc.). The digital and the physical are therefore inescapably linked in information appliances.

Prototypes are used to physically explore an idea very early on in the design process and *interactive* prototypes can be used to explore the digital considerations integrated with the physical form. However, prototypes at this initial stage are inherently 'quick and dirty': they should not require a lot of time to make and should be an exploration of an idea rather than a refined model (what Schrage [1] describes as '*Serious Play'*).

T. Gross et al. (Eds.): INTERACT 2009, Part I, LNCS 5726, pp. 217–230, 2009.

There are many academic and industry research groups working on tools and techniques for rapid interactive prototyping. These include:

- *Paper Prototyping* [2] – a very low tech approach requiring no technical skills; the user usually interacts with a paper-based version of the interface on a physical model and the screen is adjusted by a facilitator, acting as the 'computer'.
- *D.tools* [3] – a toolkit with bespoke hardware and software.
- *Phidgets* [4] and *Arduino* [5] – both provide electronic 'building blocks' to integrate into a prototype.
- *IE (Information Ergonomics) System* [6] – a flexible system of hardware and software linking a prototype to a PC.

One of the underlying recognitions that tie all this work together is that prototypes need to be made quickly in order to evaluate the tangible interactions.

The *fidelity* of a prototype is usually considered to be the resolution (the refinement and detail) of the model. A number of publications have been focused on the effect of fidelity and the advantages and disadvantages of different prototyping techniques. Sefelin et al. [7] looked at the user's willingness to criticize paper prototypes versus their willingness to criticize computer based models. Virzi et al. [8] found that there was little difference in usability data for high and low fidelity models of standard two dimensional graphical interfaces and an interactive voice response system. McCurdy et al. [9] argued for a mixed approach that allowed various aspects of a prototype to be built at different fidelity levels according to the design component being prototyped. They go on to suggest that there are five 'dimensions' or fidelity aspects that can be defined as somewhere between high and low within the same prototype, namely, aesthetics, depth of functionality, breadth of functionality, richness of data and richness of interactivity. So far this concept of mixed fidelity has been trialed with software but not physical prototypes.

Information appliances and therefore prototypes of information appliances are inherently physical. Physicality as a term, is becoming more recognized with two International Workshops on Physicality [10, 11] held recently, plus Don Norman's article on Physicality [12]. Physicality is loosely understood as being the physical nature of something, for example, a form, process or button.

This paper seeks to contribute to our understanding of the nature of physicality in the design of information appliances so that designers can become more aware of when and how to use it. To this end, we explore physicality in the context of fidelity through user trials conducted on a conceptual information appliance.

2 Background

Gill et al. [13] conducted a number of trials on a wireless home phone. They demonstrated that low fidelity physical prototypes can produce similar usability results as the end product, thus significantly outperforming touch screen mock-ups. They went on to test prototypes of decreasing fidelity until they reached a point where the similarity of user test results started to differ significantly from the results produced from the real product. They concluded that if prototypes compromise on the physical attributes of a design, such as removing the tactile feedback of the buttons, then the performance data was affected. They state that "it is not the level of fidelity that is important but rather the considerations of tangibility and physicality".

Lim et al. [14] conducted trials on a mobile phone in order to understand the effect of fidelity levels on usability data. Three prototypes were tested: the final device, a software representation and a paper prototype. All models picked up major usability issues, but only the final device and software models facilitated the collection of comments regarding the concept's comparison with other products and performance.

In our study, we interpret user data from a trial of a conceptual device as there is no completed device to compare it to. The considerations that have driven the fidelity level and its effects on the physicality of the model have been purely time based. The designer had to decide on the best way to prototype the technical aspects within the allocated time.

User trials were chosen as a means of exploring the effects of fidelity and the resultant physicality on the prototypes. The research of Gill et al. [13] and Lim et al. [14] demonstrate that user trials are an effective way of highlighting design issues by comparing low fidelity models with the final design. Those results gave us the confidence to use similar trials on a conceptual device where there is no 'end product' to compare it to. The aim of comparing the prototypes in this manner was to gather data that enabled a review of the differences in the *way* the prototypes function as each of the prototypes has the same *level* of functionality.

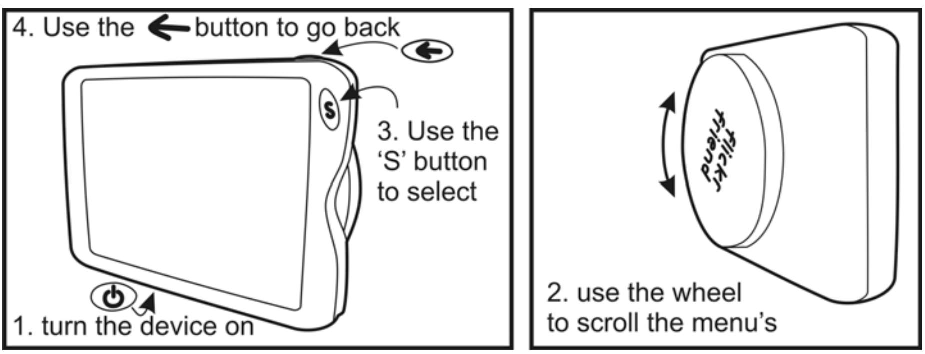


Fig. 1. Different ways of interacting with the device

3 Our Approach

The trials were conducted on a conceptual device. None of the users had been exposed to the device previously. The concept originated from an undergraduate design brief and was based on the design of a hard drive equipped device offering users the ability to wirelessly view their Flickr [15] web pages and store photos. Flickr is an online photo management and sharing application.

Some initial design work was undertaken in order to develop the physical and digital components of the concept, in order to reach a stage where, in a real design process, an interactive prototype would be the next natural step (see Figure 1 which shows the different ways of interacting with the concept). Each of the resulting prototypes used this initial design work as the starting point, therefore only the time to construct the prototype differed.

3.1 The Resulting Prototypes

'Lowest Level': Time allowed = 4 hours (actual time taken = 3 hours 30 minutes)
Method used: Paper prototyping

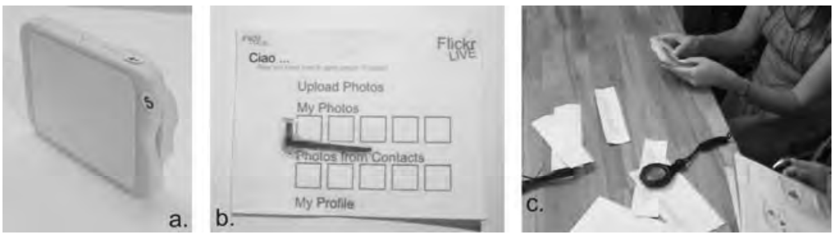


Fig. 2. Lowest level prototype: a) the foam model, b) a paper screen c) the trial set up

As noted earlier, Paper Prototyping is a very simple technique which provides a very fast method for creating low fidelity prototypes. A foam model was constructed to create the physical form to scale. The foam was sanded to produce a smooth finish with white cardboard depicting the buttons and screen (Fig. 2a). For the digital aspect a series of paper screens were created with a small red box to indicate which menu item is active (Fig. 2b). The participant held the physical model, the facilitator changed the screens and adjusted the 'select box' during user trials (Fig. 2c).

'Mid Level': Time allowed = 14 hours (actual time taken = 12 hours)
Method used: IE System

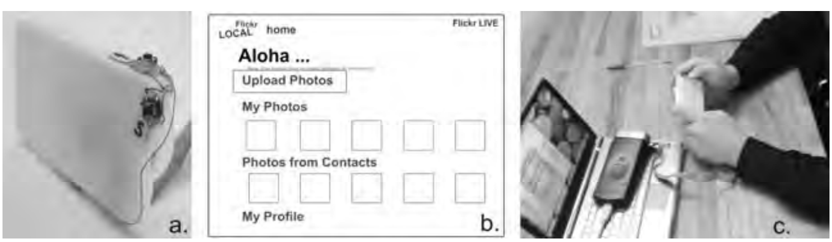


Fig. 3. Mid level prototype: a) the FDM model b) the basic flash interface c) the trial set up

The IE System was chosen to create the mid level prototype due to the simplicity it offered. The system allows a PC to receive keyboard inputs so that when a user interacts with a switch in the physical model, the PC will respond to the perceived keyboard input and a keyboard-triggered GUI is activated on the PC. A model was created in a Computer Aided Design (CAD) system and was constructed to scale using a Fused Deposition Modeling (FDM) machine (Fig. 3a). FDM is a rapid prototyping technique where the machine builds the material up layer by layer. A basic menu structure was created in Adobe Flash (Fig. 3b). The Flash animation used keyboard presses activated by off-the-shelf buttons for the screen changes, these were crudely tacked onto the outside of the model and a mechanical rotary dial was glued inside the model for the ‘wheel’ interaction. For the trial, the physical model was connected, with a cable, to a PC via the IE Unit (Fig. 3c) and the visual feedback was on a desktop monitor.

‘Highest Level’: Time allowed = 5 days (actual time taken = 5 days)
Method used: IE system and Phidgets

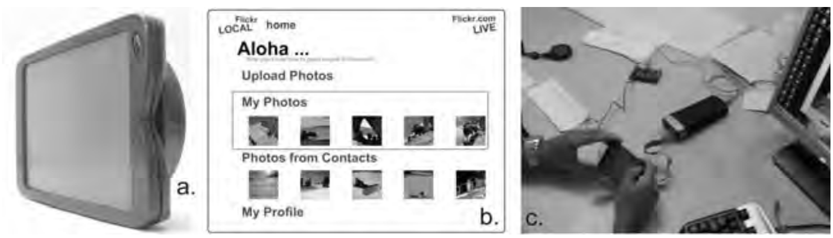


Fig. 4. Highest level prototype: a) the sprayed FDM model b) the flash interface c) the trial set up

The extra time allowed for the highest level prototype was used to develop the following three areas: the prototype was given a realistic finish, the wheel interaction was made to feel smooth and the Flash animation was developed to operate more like the intended design. Again a CAD model was created with design details such as shaped buttons and ports included. Once the FDM model had been made it was sanded and sprayed (Fig. 4a). Dome switches that produce positive tactile feedback with a low profile were used for the buttons triggering the Flash animation through the IE Unit. The smooth feeling analogue dial was an off-the-shelf Phidget component. This reflected the intended physical-digital interaction of the design intent better than the rotary dial used in the mid level prototype. The Flash animation had more realistic menus and a smoother transition between screens (Fig. 4b). For this trial, the physical model needed to be connected through both an IE Unit and a Phidget Interface Kit with wires (Fig. 4c), and the visual feedback was on a desktop monitor.

3.2 Initial Analysis of the Prototypes Created

The resulting prototypes differed considerably and their properties are reviewed in relation to McCurdy et al. [9] five dimensions of fidelity, as shown in Table 1. A similar technique is applied in Table 2 to analyse the subsequent effects on physicality, which are considered to fall under two areas: the physicality of the device itself (e.g. form, finish, weight) and the physicality of the interaction (feel of the buttons and wheel in this case).

Table 1. Properties of each prototype in relation to the five dimensions of fidelity (McCurdy et al. [9])

Dimension of fidelity	Driving factors	Lowest level 3 hrs 30min	Mid level 12 hours	Highest level 5 days
Aesthetics	Model material Model finish	**Blue foam** (both material and finish differ considerably from intended design)	**Unfinished FDM** (similar material but finish differs considerably from intended design)	**Sanded & sprayed FDM** (similar material and finish to intended design)
Richness of interactivity	Wheel mechanism	**Free rotating** (similar to intended design but no real-time feedback given)	**'Clunky', clicking mechanism with end points** (very different from intended design but gives real-time feedback)	**Smooth mechanism with end points** (very similar to intended design and gives real-time feedback)
	Buttons	**Cardboard representations** (very different in feel and aesthetics from intended design)	**Switches tacked onto model** (very different to intended design but gives real-time feedback)	**Integrated switches** (very similar to intended design in look and feel gives real-time feedback)
	Screen operation	**Paper screens** (no real-time feedback so very different from intended design)	**Basic Flash animation** (real-time feedback but sketchy interface, differs slightly from intended design)	**More advanced Flash animation** (real-time feedback and graphics similar to intended design)
Depth of functionality	Screen operation	All have identical features enabled, feature will appear 'unavailable' if it is not part of a task		
Breadth of Functionality	Screen operation	All have identical menu structures, the tasks chosen highlighted the breadth of functionality in the intended design		
Richness of Data	Data used	**Sketch data used** (different from intended design)	**Sketch data used** (different from intended design)	**Photos used** (very similar to intended design)

Table 2. Properties of each prototype in relation to the areas of physicality

Area of Physicality	Driving factors	Lowest level 3hrs 30min	Mid level 12 hours	Highest level 5 days
Physicality of the device	Scale Model material Screen material Weight	**1:1, made from blue foam with a cardboard screen** (form is very similar to intended design, finish and weight is considerably different)	**1:1, unfinished FDM with screen placement suggested on model (no colour difference)** (form is very similar to intended design, weight and finish are considerably different)	**1:1, finished and sprayed FDM with a colour difference depicting the screen** (form and surface finish is very similar to intended design, weight is different)
Physicality of the interaction	Wheel mechanism	**Wheel freely rotates** (as intended in design) **with no real-time physical or digital feedback** (extremely different from intended design)	**Mechanism feels clunky and cannot rotate continuously** (considerably different from intended design) **gives real-time physical** (not part of intended design) **and digital feedback** (part of intended design)	**Mechanism feels smooth** (very similar to intended design), **cannot rotate continuously** (not part of intended design) **gives real-time physical and digital feedback** (similar to intended design)
	Buttons	**Buttons are depicted with cardboard and give no physical or digital feedback** (very different to intended design)	**Buttons are off-the-shelf and tacked onto the model** (very different to intended design) **but give real-time physical and digital feedback** (similar to intended design).	**Buttons are integrated dome switches with real-time digital and physical feedback** (very similar to intended design)

4 Method

The set of trials and rating scale used to classify the severity of problems, was based on recommendations by Redish et al. [16]. Participants were divided into three independent groups, with each group using one level of prototype (low, mid or high). Each participant was given a series of 5 scripted tasks [17]:

Task 1: turn the device on
Task 2: find a photo on the *Flickr* website
Task 3: find a friend photo on the *Flickr* website
Task 4: find a photo from the hard-drive
Task 5: transfer a photo from a camera

4.1 Structure of the Trials

The following structure was applied to every participant for each of the three prototypes trialed:

i. Participant fills in a demographic questionnaire covering age and gender plus existing technology usage. Note the prototype is not in sight at this stage.
ii. Participant is given a written description of the product.
iii. Facilitator uncovers the model and records if the participant picks it up and her reaction in relation to the fidelity of the aesthetics.
iv. Participant is given the 5 tasks (as described above) to carry out. Facilitator records the time taken for each task and whether the user experienced a success, minor problem, serious problem, or a catastrophe (see Table 3).
v. Participant fills in a questionnaire and is asked to rate certain aspects of their experience with the device.

4.2 The Empirical Study

A pilot study was first carried out with 9 undergraduate participants from the University of Wales Institute, Cardiff (UWIC), which uncovered some problems, including hardware stability issues, and these were then fixed.

The main study was conducted using 48 participants recruited from UWIC staff who have used digital cameras (including cameras on their mobile phones). The participants were divided into three groups of 16, one for each fidelity level, to eliminate possible learning effects. 23 females and 25 males were trialed with ages ranging from 19 to 50, thus an average age of 29. All trials were videotaped for further qualitative analysis.

5 Quantitative Analysis

The quantitative data of interest is the 'performance' data, which shows whether the task was a success, had minor or major problems or was a catastrophe. The data was recorded at the time of each trial based on the criteria shown in Table 3.

Table 3. Description of performance rating and examples

Performance rating	Definition	Examples
Success	Task completed without error	User finds all the correct buttons and menus when needed
Minor problem	Task completed with small error	User goes into the wrong menu, user cannot find a button
Major problem	Task completed with major error/s	User repeatedly tries the wrong menus or buttons
Catastrophe	Task is not completed	User has not completed a task (even if he/she thinks they have), user gives up.

The quantitative analysis was conducted in order compare the results of the prototypes for each of the separate tasks (repeated measures). The performance data was converted into interval data (3 = success; 2 = minor problem; 1 = major problem; 0 = catastrophe) and analysis was conducted using a 3 (prototype level) by 5 (tasks) mixed analysis of variance (ANOVA) with the alpha level set to 0.05.

Figure 5 shows the performance data, a line has been included between the marks to aid interpretation of the graph. No significant overall differences were found between the prototypes. The plots suggest that the prototypes performed similarly for

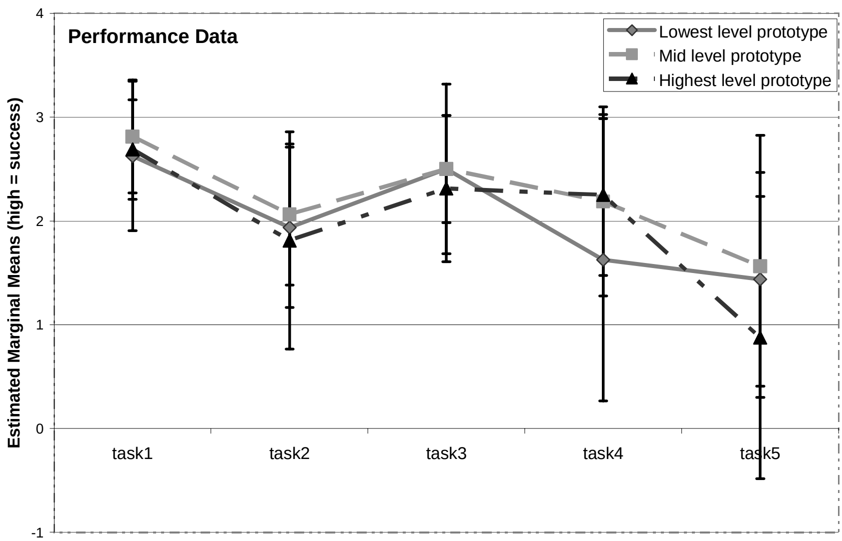


Fig. 5. Performance ratings for each of the 5 tasks as a function of device type

Tasks 1-3, but Tasks 4 & 5 appear to show some differences. Upon further analysis (simple effects) these differences were found to be not significant.

The quantitative data on its own did not reveal any differences, which suggests that neither differences in physicality nor in fidelity have an effect, or that this is not a reliable way of analyzing this effect.

6 Qualitative Analysis

The qualitative analysis was conducted by reviewing the video recordings of each participant after the trials. The qualitative analysis was twofold: firstly, identifying problems that participants may have encountered while performing each task (Part 1 Analysis) and secondly, assessing whether participants were influenced by the fidelity and physicality of the prototypes (Part 2 Analysis).

Part 1 Analysis: This was conducted to find out where participants were having problems performing each task (*types of usability problems)*. During the trials, the main errors were observed and noted in a table. Later, while reviewing the video, each error made by the participant was recorded. If an error had not been listed before, it was added to the table. However, if a participant kept repeating the same error, it was recorded several times, this highlighted particular areas of concern. The errors were then condensed into four problem areas, which we identified as being of hindrance to a user in completing a task. The problems areas are:

a. Unclear meanings of symbols
b. Difficulty locating appropriate interface elements
c. Unexpected feedback from software (mental model mismatch)
d. Unintentional interaction with software (wanted to interact in a way that was not intended)

Table 4. Number of times usability problems occurred at different prototype level

Task	Usability problems	Prototype level		
		low	**med**	**high**
1	Locating appropriate interface element	9	4	9
	Got it right first time	12	14	11
2	Unclear meanings of symbols	2	1	5
	Locating appropriate interface element	32	13	18
	Unexpected feedback from software (mental model mismatch)	2	6	18
	Unintentional interaction with software	4	0	9
	Got it right first time	2	5	6
3	Unclear meanings of symbols	2	4	1
	Locating appropriate interface element	2	2	8
	Unexpected feedback from software (mental model mismatch)	12	2	16
	Got it right first time	10	9	8
4	Locating appropriate interface element	2	0	2
	Unexpected feedback from software (mental model mismatch)	22	17	12
	Unintentional interaction with software	6	4	4
	Got it right first time	7	8	6
5	Locating appropriate interface element	0	8	3
	Unexpected feedback from software (mental model mismatch)	23	18	16
	Unintentional interaction with software	1	4	8
	Got it right first time	3	4	4

Table 4 shows the number of times users encountered usability problems for each task at different prototype levels. The results that are of particular interest are those that differ across the prototypes. So, for example, during Task 2 there were 2 problems recorded by the lowest level prototype due to unexpected feedback from the software but. The same task resulted in 18 problems for the highest level prototype. Other notable results are again for Task 2 where users could not locate the appropriate interface 32 times for the lowest level, 13 times for the mid level and 18 times for the highest level. Interestingly for Task 5, there were 0 problems for the lowest level prototype in locating the appropriate interface elements, but 8 problems for the mid level and 3 problems for the highest level.

Further analysis of the problems related to Task 2 suggests that users of the lowest level prototype had so much trouble identifying the correct interaction (32) that there were very few mental model mismatch issues (2). Compare this to the highest level prototype, where users were able to find the interaction better (18 errors), but they had difficulty with the mental model of the device (18). The inability to identify the correct interaction could arise either because of a lack of understanding of the symbols (which were the same across the prototypes) or a complete misunderstanding of the results of that form of interaction. The mid level prototype instead has the lowest number of problems related to 'identifying the interaction' (13) and an average range of problems with the mental model (6). So what could be the reason behind these problems? From Table 2, we can see that the lowest level prototype has no tactile feedback on pressing the buttons (just the facilitator moving a screen), while the mid level prototype has very pronounced buttons that give both tactile and on screen feedback, and the highest level prototype has more subtle visual properties with subtle tactile feedback plus on screen feedback. The number of problems linked with locating the appropriate interface element in Task 5 could have arisen due to the same issues as in Task 2, in other words, users of the lowest level prototype had already made so many mistakes early on that they are less likely to make mistakes in the later tasks, unlike users of the highest level prototype who are still experiencing problems even in the later tasks.

Part 2 Analysis: This was undertaken to assess whether participants were affected by the fidelity and physicality of the prototypes based on the related comments made, for example, ‘wheel mapping not natural’. A similar recording procedure was followed as in Part 1 Analysis using the errors noted during the trials plus the video review. The comments were then sorted and the ones related to the following areas were selected:

1. physicality of the device (e.g. size in the hand, screen position and size)
2. physicality of the interaction (e.g. the button is in the wrong place, how the wheel feels etc.)
3. feedback about the design and idea in general

The results are shown in Table 5. The general feedback on the design and concept is roughly the same across the prototypes. The lowest level prototype seems to differ in the number of comments about both the physicality of the *device* (22 at the lowest level compared to 13 at the mid level and 16 at the highest level) plus the physicality of the *interaction,* 42 at the lowest level compared to 52 at the mid level and 57 at the highest level. These results suggest that the test was set up in a way that entices generally attracted more comments about the physicality of the interaction rather than the physicality of the device. However, the lowest level prototype received more comments about the physicality of the device unlike the mid and highest levels. This could be because the physicality of the interaction was so far removed in the lowest level prototype from that intended, hence it was harder for users to judge this aspect of the design and as a result, they made more comments about the physicality of the device itself.

Table 5. number of comments related to the physicality and fidelity at different prototype levels

	Lowest level	Mid level	Highest level
Physicality of the device	22	13	16
Physicality of the interaction	42	52	57
General feedback on the design and	19	17	18

7 Discussion

Each of the prototypes created represented the same design intent and enabled the same functionality. Time constraints governed the fidelity level and each prototype was tested for usability and physicality issues. The prototypes needed to convey enough information to the users so they were able to get a feel for the design intent of the product. The initial hypothesis was that fidelity and subsequently physicality would have an effect on the users understanding of the product and therefore user feedback and usability would be affected. The analysis of the user trials showed the following results:

1. All prototypes achieved similar results for the performance test.

There was in fact little difference in performance across the prototypes with different fidelity levels (which would seem to agree with the research by Lim et al. [14]). This in itself is an important result showing that in the early stages of the design process, the fidelity level might not have a significant impact. Despite the mid level prototype being physically different from the intended design in a number of seemingly important ways (the wheel clicked, could not rotate 360° and felt very 'clunky'), it still produced valid feedback about the concept. Furthermore, the mid level prototype took less than half the time to build compared to the highest level prototype. Even the lowest level paper prototype seemed to produce usability data in line with the higher fidelity ones.

2. Users of the mid and highest level prototypes, with real time tactile and digital (on screen) feedback, had fewer problems in locating the appropriate interface element.

Even when all the prototypes had the same graphical symbols, the lowest level prototype users had a lot of problems identifying the appropriate interface elements. This may be because many users worked out what interactions did by 'experimenting' with them instead of understanding the symbols and thus made their decisions based on the feedback they received. This approach was only supported by those prototypes that had real time feedback, whereas the lowest level paper prototype required the facilitator to find the correct feedback and change the screen.

3. Users of the mid and highest level prototypes had more problems with the mental model of the device early on in the trial whereas the lowest level prototype users encountered these issues later on in the trial.

This is an unexpected outcome. Table 4 shows that, even after completing 4 tasks, users still made errors due to a mental model mismatch for task 5. Users who had real time tactile and digital feedback from their interactions had more difficulty in understanding how the device worked. The most likely explanation for this is that users of the lowest level prototype were so distracted by not locating the appropriate interface element that this overshadowed their understanding of the device. By the end of the trial, users of the lowest level prototype were having less problems locating the interface element but more difficulty in understanding the device (their mental model). This could possibly be due to the users' inability to fully engage with the device and therefore following a 'more luck than judgment' approach.

4. The mid and highest level prototypes gave more feedback about the physicality of the interaction.

This was not unexpected as in order to get valid feedback about an interaction, one needs to approximate the intended interaction, which the lowest level paper prototype did not facilitate.

8 Conclusion and Further Work

This paper has reported on a preliminary investigation into the effects of physicality and fidelity on the prototypes used for front end product design development. The trials suggest that there is no effect of fidelity at the early stage of the design process in terms of user performance, however a deeper analysis is required. As expected, the qualitative analysis showed that prototypes that gave real time interaction and feedback allowed users to get a more realistic appreciation of how the device worked, and also generated more useful comments about how the device feels to hold and to interact with.

From these results, we can draw that for the initial exploration of a design idea, very low fidelity prototyping is a fast and low cost method of getting reliable feedback. On the other hand, if more specific feedback about the intended design and interaction is required, then a prototype that can produce immediate feedback is essential. However, there are many more factors at play and these need to be researched further to inform design guidelines in relation to the needs of the early design process.

The nature of physicality seems to have an impact on the user trials of these prototypes, but a very in-depth analysis had to be carried out to tease out these effects. It would be more useful if such effects could be found and explored using faster quantitative analysis. Further work needs to be undertaken to explore how these effects of physicality can be tested in a quantifiable way and therefore fully explore the wider implications for designers in practice building.

Acknowledgments. This work has been supported by the AHRC/EPSRC funded DEPtH “Designing for Physicality” (www.physicality.org) project which is part of the “Designing for the 21st Century” research initiative (http://www.design21.dundee.ac.uk/).

References

1. Schrage, M.: Serious Play: How the World’s Best Companies Simulate to Innovate. Harvard Business School Press, Boston (1999)
2. Snyder, C.: Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces. Morgan Kaufmann, San Francisco (2003)
3. Hartmann, B., Klemmer, S.R., Bernstein, M., Abdulla, L., Burr, B., Robinson-Mosher, A., Gee, J.: Reflective physical prototyping through integrated design, test, and analysis. In: Proceedings of the 19th annual ACM symposium on User interface software and technology. ACM, Montreux (2006)
4. Greenberg, S., Fitchett, C.: Phidgets: easy development of physical interfaces through physical widgets. In: Proceedings of the 14th annual ACM symposium on User interface software and technology. ACM, Orlando (2001)
5. Burleson, W., Jensen, C.N., Raaschou, T., Frohold, S.: Sprock-it: a physically interactive play system. In: Proceedings of the 6th international conference on Interaction design and children. ACM, Aalborg (2007)
6. Gill, S.: Developing Information Appliance Design Tools for Designers. In: 1st Appliance Design Conference, Bristol, UK (2003)

7. Sefelin, R., Tscheligi, M., Giller, V.: Paper prototyping - what is it good for?: a comparison of paper- and computer-based low-fidelity prototyping. In: CHI 2003 extended abstracts on Human factors in computing systems. ACM, Ft. Lauderdale (2003)
8. Virzi, R.A., Sokolov, J.L., Karis, D.: Usability problem identification using both low- and high-fidelity prototypes. In: Proceedings of the SIGCHI conference on Human factors in computing systems: common ground. ACM, Vancouver (1996)
9. McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B., Vera, A.: Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success. In: Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM, Montreal (2006)
10. Ghazali, M., Ramduny-Ellis, D., Hornecker, E., Dix, A. (eds.): Proceedings of the First International Workshop on Physicality, Physicality 2006, Lancaster University, UK, February 6-7 (2006)
11. Ramduny-Ellis, D., Dix, A., Hare, J., Gill, S.: Proceedings of the Second International Workshop on Physicality, Physicality 2007, Lancaster University, UK, September 2-3. UWIC Press (2007)
12. Norman, D.A.: The next UI breakthrough, part 2: physicality, vol. 14, pp. 46–47. ACM, New York (2007)
13. Gill, S., Walker, D., Loudon, G., Dix, A., Woolley, A., Ramduny-Ellis, D., Hare, J.: Rapid Development of Tangible Interactive Appliance: Achieving the Fidelity/Time Balance. International Journal of Arts and Technology (IJART) Special Issue on Tangible and Embedded Interaction 1, 309–331 (2008)
14. Lim, Y.-k., Pangam, A., Periyasami, S., Aneja, S.: Comparative analysis of high- and low-fidelity prototypes for more valid usability evaluations of mobile devices. In: Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. ACM, Oslo (2006)
15. Flickr (2004), `http://www.flickr.com`
16. Redish, J., Bias, R.G., Bailey, R., Molich, R., Dumas, J., Spool, J.M.: Usability in practice: formative usability evaluations - evolution and revolution. In: CHI 2002 extended abstracts on Human factors in computing systems. ACM, Minneapolis (2002)
17. Rubin, J.: Handbook of Usability Testing. John Wiley & Sons, Chichester (1994)

Appendix 6: Paper work for Study One

Project Summary:

You have been asked to participate in the user trial of the concept 'Flickr Friend'.

We are undertaking this research as part of a wider research project within UWIC that seeks to explore the physical nature products and prototypes in relation to user experience.

Why You Have Been Asked to Participate:

In order for us to test the product (Flickr Friend) fully we need participants who have had no previous involvement with the product.

Your participation is entirely voluntary.

It is important to understand that we are testing the design of the product, not the user, if there are any difficulties please explain them so we can eliminate them from the design.

During the test we will give you five tasks to perform on the product and then ask for feedback from you via a questionnaire that we will complete after the tasks.

The complete user trial should last around 30mins.

Project Risks

The research involves a questionnaire-based interview which will be videoed for later analysis, and we are not seeking to collect any sensitive data on you. We do not think that there are any significant risks associated with this study. If you do feel that any of the questions are inappropriate then you can stop at any time. Furthermore, you can change your mind and withdraw from the study at any time – we will completely respect your decision.

How We Protect Your Privacy

All the information we get from you is strictly confidential. We have taken careful steps to make sure that you cannot be directly identified from any of the questionnaire forms. We will keep your name and any personal details completely separate from the other questionnaire forms, and there is no information on these questionnaires that will easily identify you. Your personal details and your finished questionnaire will be kept are secure locations within UWIC. When we have finished the study and analysed all the information, all the forms used to gather the data will be destroyed.

<table>
<tr><td colspan="8">User Number:</td></tr>
<tr><td colspan="8">Age:</td></tr>
<tr><td colspan="8">Gender:</td></tr>
<tr><td colspan="8">What mobile phone do you currently have?</td></tr>
<tr><td colspan="8"></td></tr>
<tr><td colspan="8">Do you use a digital camera?</td></tr>
<tr><td>No, don’t own one</td><td></td><td>Only use it rarely</td><td></td><td>Take it only on holiday, trips, special occasions</td><td></td><td>Carry it around most of the time</td><td></td></tr>
<tr><td colspan="8">Do you use the camera on your mobile?</td></tr>
<tr><td>Don’t have one on my phone</td><td></td><td>Rarely use the one on my phone</td><td></td><td>Use it only when I haven’t got my dedicated camera</td><td></td><td>Use it all the time</td><td></td></tr>
<tr><td colspan="8">Do you have a Flickr account?</td></tr>
<tr><td>No</td><td></td><td colspan="2">Yes, the free account</td><td></td><td colspan="2">Yes, the pro account</td><td></td></tr>
<tr><td colspan="8">Do you use any other photo sharing or management system? If so, what?</td></tr>
<tr><td colspan="8"></td></tr>
<tr><td colspan="8">How do you currently store your photographs?</td></tr>
<tr><td colspan="8"></td></tr>
<tr><td colspan="8">What mobile products do you normally carry with you? Please include brand name where possible. Eg MP3 player, PDA….</td></tr>
<tr><td colspan="8"></td></tr>
</table>

The Effect of Physicality on Low Fidelity Interactive Prototyping for Design Practice

Joanna Hare, Steve Gill, Gareth Loudon, and Alan Lewis

Cardiff Metropolitan University, Cardiff, UK
jhare@pdronline.co.uk,
{sjgill,gloudon}@cardiffmet.ac.uk,
profalan.lewis@wales.ac.uk

Abstract. In this paper we propose the concept of 'active' and 'passive' physicality as mental models to help in understanding the role of low fidelity prototypes in the design process for computer embedded products. We define ‘active physicality’ as how the prototype and its software react to users and ‘passive physicality’ as how the prototype looks and feels offline. User trials of four different types of ‘low fidelity’ prototypes were undertaken using an existing product as the datum. Each prototype was analysed in terms of active and passive physicality and user responses were collated and compared qualitatively and quantitatively. The results suggest that prototypes that balance both active and passive physicality produce data closer to the final device than those that are strong in one at the expense of the other.

Keywords: Physicality, interactive prototypes, computer embedded products, design, product design, iterative product development, information appliances.

1 Introduction

This paper builds on previous research on physicality and low fidelity interactive prototypes. Virzi et al. [1] found that there was little difference in usability data for high and low fidelity models of standard two dimensional graphical interfaces and an interactive voice response system. Yet a number of researchers [2] [3] felt that the concept of low verses high fidelity is not quite enough to convey the whole manner of situations that prototypes are constructed for. McCurdy et al. [3] argued for a mixed approach that allowed various aspects of a prototype to be built at different fidelity levels according to the design component being prototyped. They go on to suggest that there are five ‘dimensions’ or fidelity aspects that can be defined as somewhere between high and low within the same prototype, namely, aesthetics, depth of functionality, breadth of functionality, richness of data and richness of interactivity. So far this concept of mixed fidelity has been trialled with software but not physical prototypes. Despite several authors conducting studies on prototypes of computer embedded devices the physical properties of both the model and interaction have been largely ignored.

In 2008 we demonstrated that in order to trial an interactive device with users an interactive prototype must be constructed [4]. The same study went on to lower both

P. Kotzé et al. (Eds.): INTERACT 2013, Part I, LNCS 8117, pp. 495–510, 2013.

the level of physical fidelity of the model and the visual fidelity of interface until usability data started to significantly differ from the results of the final device. It was proposed that subtle differences in physicality, in this case removing the tactile feedback of buttons, affected the results suggesting that considerations of physicality are more important than the level of fidelity. This poses the question of how we 'consider' physicality.

However, a study published in 2009, demonstrated that some effects of physicality on user trials were only apparent through in-depth analysis because the effects were often subtle and the picture sometimes confusing [5]. This study seeks to clarify the position physicality occupies in user interactions.

The 2009 study sought to uncover the resulting differences in physicality based on low, medium and high(er) fidelity prototypes. In this study physicality was considered to fall under two areas: the physicality of the device (e.g. form, finish, weight) and the physicality of the interaction (the feel of the buttons and wheel in this case). But this method only allows the prototypes to be described and not directly compared which is essential when using physicality to determine the differences between the prototypes on trial. The physicality of the device and interaction was an appropriate way to describe the prototypes and, with subsequent analysis, this has been adapted to form the concept of 'passive' and 'active' physicality where:

Passive Physicality is how the prototype looks and feels when turned off, for example the weight, finish and button locations.

Active Physicality is how the prototype reacts to the users, typically the reaction of the interface (software), the feel of the buttons when operated (or sliders, dials, screen etc.)

To explain these terms a useful starting point is that of Dix et al. [6] who regard the physical device removed from its context and 'separated' from its digital operation in order to consider the mapping of the device 'unplugged'. This is the basis of 'passive' physicality; the judgments that can be made about the device without switching it on. Do you grasp a cup by its handle or by the body? Decisions are made about the comfort of the cup's handle by its appearance and the perceived weight of the contents of the cup [7]. Passive physicality also has its roots in Gibson's description of affordances [8] which suggest ways of interaction. Affordances are not simply a property of the object; they are the way a specific user relates to that object. When Norman [9] applied Gibson's idea to design; he divided the idea of affordances into those of real and perceived affordances. Whilst real affordances tell the user what they could actually do with the device, meaningful or not, perceived affordances tell the user 'what actions can be performed on an object and, to some extent, how to do them'. Yet passive physicality is more than affordances, it includes the physical properties of the device, its weight, finish and locations of the interactions.

Active physicality is concerned with the interactive portion of the device; what happens when the device is being used. It is still the physical that is of concern but in relation to the device's purpose and ease of use; how buttons operate the interface and how those buttons (or any interactions) feel when operated.

The exact drivers behind active and passive physicality might differ depending on the product being prototyped but the essence of active and passive physicality will remain.

This study proposes that a prototype can be considered by its level of active and passive physicality. For example, a prototype that is driven by the technology of the experience rather than the proposed size of the design would have a high level of active physicality but low passive physicality.

By attempting to understand physicality and using this to drive the physicality of low fidelity prototypes we aim to draw out just how physicality can be used by the designer to create efficient low fidelity prototypes. The efficiency of a prototype is of great importance; an efficient prototype can supply reliable data for a fraction of the cost of a high fidelity prototype enabling an iterative process. The early stages of the typical user-centred design process are highly iterative in order to react to and inform the developing project. User trials are a key tool to gathering data needed to inform the project, techniques include rapid ethnography [10], usability evaluation [11] and task centered walkthroughs all of which can be supported by interactive prototypes, and these prototypes need to be fast, low-cost and stage appropriate. This paper presents an early stage study on four low fidelity prototypes of the same device.

2 Methodology

2.1 The Prototypes

An existing product was chosen to provide a datum against which the retrospectively developed prototypes could be measured. The choice to retro-prototype an existing device as a method was taken after considerable thought. The alternative would have been the development of a new device. Both methods have been used in prototype evaluation studies [4] [12]. Retro-prototyping was chosen because it has the benefit of access to a real, mass produced product, identified by the manufacturer as a worth-while idea and having successfully undergone a product development process. The finished device can be used to compare the results from the user study in a manner that is all but impossible to recreate in a research study.

Fig. 1. The iRiver SPINN

The product chosen was the iRiver Spinn (Figure 1), a personal music player. The main features and interactions of the iRiver Spinn are shown in Figure 2.

Four low fidelity prototypes were constructed using techniques currently in use in industry. Each prototype was planned giving due consideration to active and passive physicality levels, with the intention of placing one in each of the quadrants shown in the graph in Figure 3.

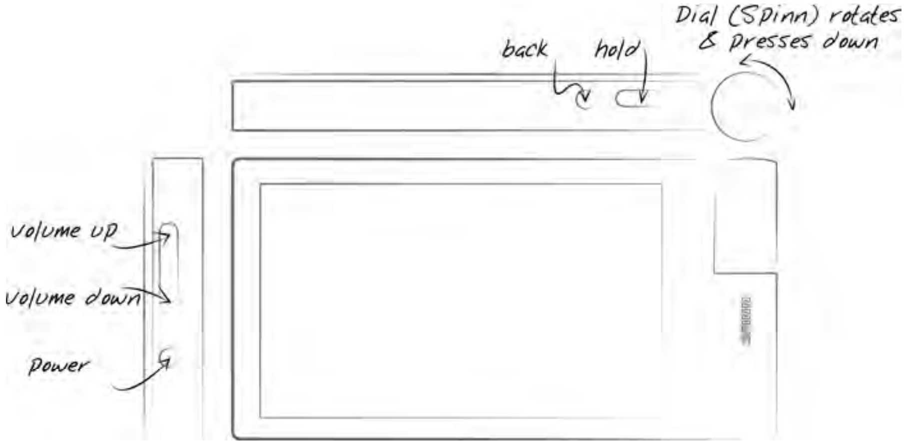


Fig. 2. The interactions of the iRiver Spinn

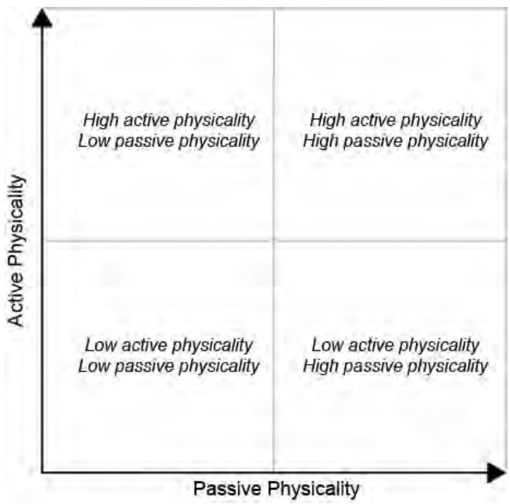


Fig. 3. Areas of physicality

The time taken to make each type of prototype is a critical issue. ‘Time is money’ and so we timed the building process and applied an hourly rate of £40 in order to cost each prototype. These are shown in Figure 4.

Fig. 4. The time taken to construct the prototypes

A single interface core was coded in Adobe Flash for all prototypes and adapted to the needs of each. Preparatory work ensured that this interface would be suitable for all prototypes and that the adaptation of the interface was possible for all. As is typical at this stage of the design process, only a limited selection of features were included in the software [11]. A single Computer Aided Design (CAD) model was created.

Prototype 1 (Figure 5; named ‘blue foam’) was constructed from model making foam board. Interaction was based on the Wizard of Oz technique [13], the Flash interface was operated remotely by the facilitator and viewed on the Tablet, the participant was asked to follow the ‘think out loud’ protocol [14], the facilitator could react to what the participant was saying and interacting with on the foam prototype.

The physical model for **Prototype 2** (Figure 6; named ‘IE4’) was constructed using rapid prototyping techniques (FDM). The CAD model was adjusted slightly to house the buttons and the dial which were integrated to make the prototype interactive. An IE4 [15][1] was used to connect the buttons to a laptop. The Flash interface, shown on a tablet, ‘listens’ for key presses from the IE4 and triggers changes in the interface when the participant interacts with the prototype.

[1] The IE4 is a wireless device which converts buttons presses into keyboard presses.

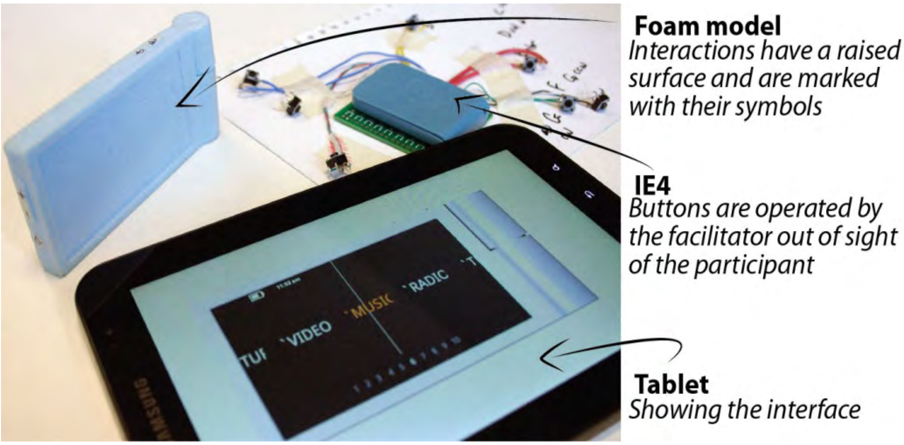


Fig. 5. Prototype 1: Foam prototype

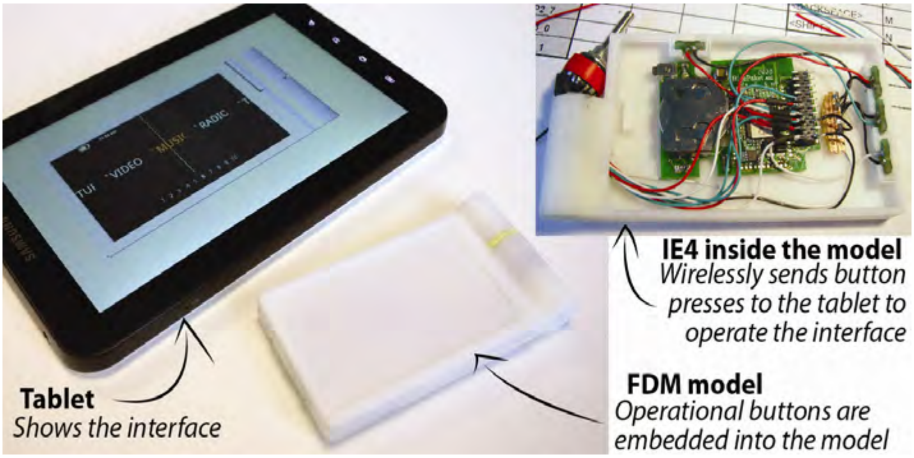


Fig. 6. Prototype 2: IE4

The physical model for **Prototype 3** (Figure 7; named 'appearance model') was intended to reflect the final device as accurately as possible. The form was rapid prototyped (using FDM) then finished to facsimile level. The Flash interface was operated by the participant on a touch screen tablet.

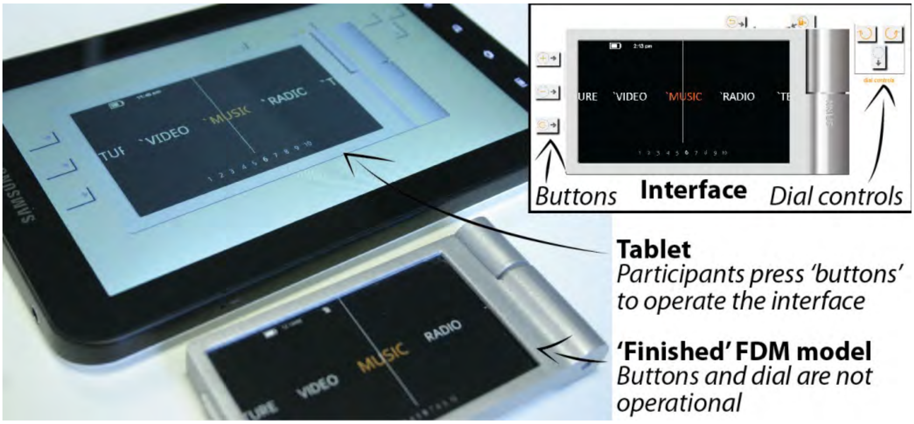


Fig. 7. Prototype 3: Appearance prototype

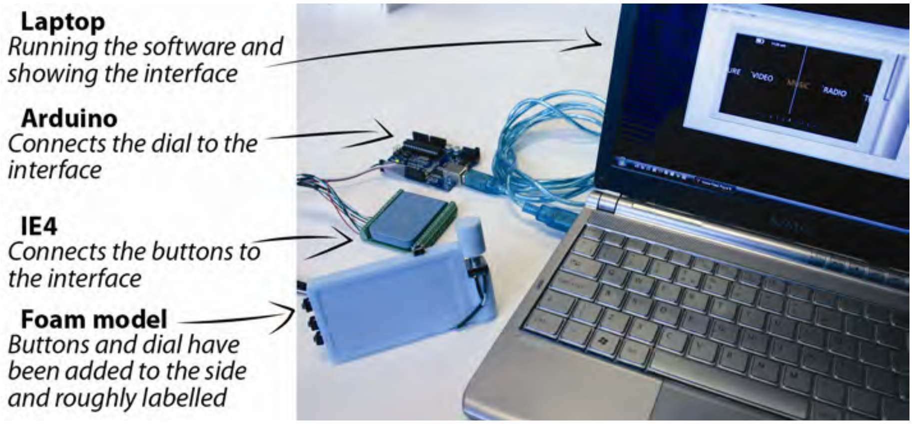


Fig. 8. Prototype 4: Arduino prototype

A rough foam model was constructed for **Prototype 4** (Figure 8; named 'Arduino') to accommodate the off-the-shelf buttons and dial. The dial was connected to an Arduino [16] which received the analogue signals and outputted them to the computer running the Flash interface. The buttons were connected to an IE4. Due to the extra code required for the Arduino, the interface was shown on a laptop rather than the touch screen tablet.

2.2 Assessing Physicality

Each of the prototypes was analyzed in terms of active and passive physicality. The main factors in the design that would determine the passive physicality levels of the prototype were determined to be: scale, form, finish and button location. For active physicality the main issues were: Spinn physical feedback, Spinn digital feedback, button physical feedback and button digital feedback. Initially a 'scoring' system was trialed but this was discarded, for when we call a prototype 'low' fidelity we do not assign that 'lowness' a value, as designers we intrinsically know when a prototype is low fidelity. It is only when conducting studies such as this that a prototype is considered lower or higher than another. Figure 9 shows the considerations for assessing each prototype.

Prototype	Passive physicality	Active physicality
Blue Foam	Low This prototype looks approximate and feels light, buttons are obviously cardboard and not working.	Low Buttons are obviously intangible and the participant is speaking through their expected interactions which are being interpreted by the facilitator who is operating the Flash based interface.
IE 4	Mid This prototype looks reasonable with no distracting wires. The prototype can be held comfortably yet it is very obviously an early stage prototype.	Mid Interactions mimic the design intent satisfactorily directly operating the interface which is a reasonable approximation of the design intent.
Appearance model	High The prototype looks and feels very similar to the final product.	Low The interactions are not obvious as the participant does not use the tangible prototype to operate the interface; instead the interface is operated on a touch screen breaking the link between the tangible product and its interface.
Arduino	Low The prototype has tacked on switches and wires are distractingly apparent in both the aesthetics and tangibility of this prototype.	High The prototype accurately mimics the way the final device feels when it is operated, both in the way the buttons work and the functionality of the interface.

Fig. 9. Assessing the levels of active and passive physicality of the prototypes

The Appearance and Arduino prototypes are high in one area of physicality at the expense of the other, whilst the Foam and IE4 prototypes 'balance' both active and passive physicality, as shown in Figure 10.

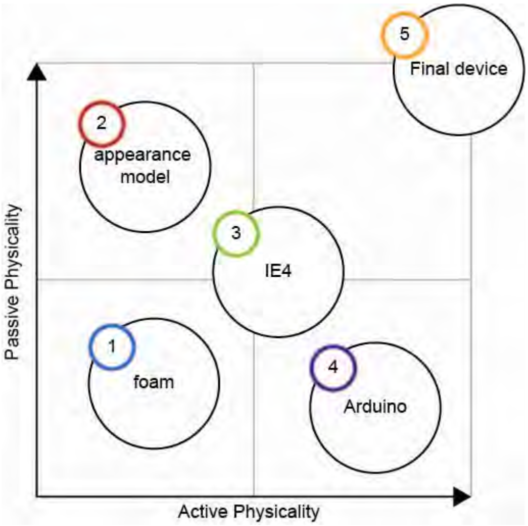


Fig. 10. The resulting physicality of each of the prototypes

2.3 The User Study

40 participants were recruited for the study (eight per prototype [17]), two did not turn up and three tests were rejected due to technical difficulties so the total number included in this analysis is 35.

16 of the participants were female and 19 were male. Participants were screened in accordance with the target market identified by iRiver to be between 23 and 45 years old; recruited participants fell predominately into the <28 (49%) or 29-33 (34%) age groups. All listened to music on a dedicated player or mobile phone and none had used the iRiver Spinn before.

Task-orientated trials, typical of usability trials, can be an effective way to demonstrate the product to a participant in a controlled manner and the participants were encouraged to 'think aloud' during the study to communicate their thought process [18]. Five tasks were chosen to introduce the participant sequentially to the device and no time constraint was imposed for the tasks. The tasks were:

Task 1: Turn the device on
Task 2: Find and play a specific track
Task 3: Adjust the volume of the track
Task 4: Stop the track and navigate to the first screen
Task 5: Turn the device off

Next, each participant was asked to scroll through the main menu titles and discuss what they expected within each menu. This user-led exploration ensured each participant had the same knowledge of the features of the device. After which a semi-structured interview sought to gain feedback about both the physical design and the users' interaction experience of the product. The explicit nature of the tasks and user-led exploration is one of the recommendations to reduce the evaluator effect on studies [19].

Finally, users were introduced to all the prototypes and asked to fill in a questionnaire ranking the quality of feel, appearance and quality of interaction for each of the prototypes. This enabled the participants to directly compare prototypes and offer an opinion about their construction.

Participants were bought into a controlled environment and the entire user trial was recorded on video. A facilitator ran the study with an observer monitoring the study via the video link. The observer was able to ensure continuity across the studies; this was deemed more suitable than introducing them as a second evaluator due to their level of experience with the prototypes and user testing methodologies. The Facilitator has conducted a number of similar studies before in a research and commercial context and is therefore able to reflect on techniques with colleagues of similar experience. Thus although the evaluator effect cannot be eliminated, it has been considered for this study [19].

3 Results of the User Trial

The analysis was performed by the facilitator. Discourse analysis provided a framework to analyse the video footage of the tasks, menu exploration and semi-structured interview. The strength of this approach is that it gives the ability to structure the conversational feedback typical of this type of study in a rigorous manner. The video footage was reviewed with event logging software and comments were assigned 'codes' based on the type of comment. 50 comment groups were recorded in total. In order to compare the prototypes comments made by just one participant were removed. These comments were then reviewed and collated to form high-level design recommendations typical of a report from user trials [20]. Further recommendations could be drawn from the data produced by the studies that would be used in a commercial context. For the purpose of this study only the comments that have emerged through the formal discourse analysis are included. It is important to note that the recommendations themselves are not important to this study and have therefore been simplified for this paper; it is the number of recommendations identified for each prototype in relation to the final device that is of importance in this context. The ten key comments that the design recommendations address are:

1. Help required from the facilitator
2. Difficulties in finding the required interaction
3. Tried other interactions
4. Pressed back to stop track playing
5. Tried turning dial to get to pause icon
6. Observation that it looks like a touch screen device
7. Like the 'Spinn' interaction
8. Long-winded interface
9. No unique selling point
10. Vertical menu navigation not obvious

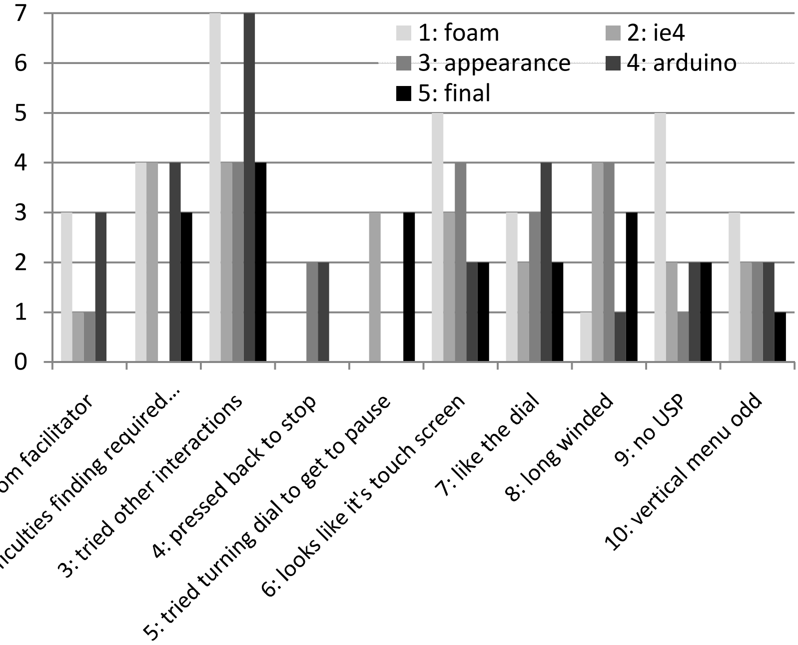


Fig. 11. The ten key comments addressed by the design recommendations

Figure 12 shows the results of the ranking exercise where each of the participants were introduced to all the prototypes and asked to give a rating where 6 is positive and 1 is negative. The participants were asked to rate three elements of the prototypes; the ‘quality of feel’ and ‘appearance’ which aimed to prompt the participant to consider the passive physicality elements and the ‘quality of interaction’ roughly equates to active physicality. Although these terms cannot be directly described as active and passive physicality, it goes some way to enable a comparison to the assessment of physicality shown in Figure 10. The data from the prototype the participant used for the study was not included to eliminate any bias from familiarity with the prototype. Figure 12 shows participants consider the foam prototype to have a low ranking but roughly equal for both elements which supports our assessment of the prototype to be low in both active and passive physicality. Likewise the appearance and Arduino are ranked in a similar way to our assessment. The IE4 gives interesting results with it being considered a higher quality of interaction than the Arduino and a more marked difference between active and passive physicality than anticipated. It could be that the visual aspects of physicality are undervalued in the current definition of passive physicality or that these questions are not adequate at obtaining participants views of active and passive physicality, this is beyond the scope of this paper but could be an interesting topic for further research. This exercise enabled participants to reflect on the prototypes themselves during the ranking exercise and the comments made were also captured, these will be brought into the discussion.

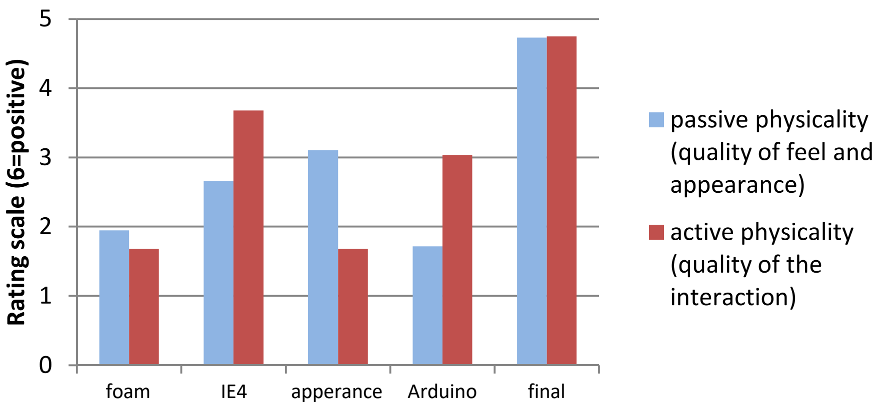


Fig. 12. Data from the ranking exercise; comparing the prototypes

4 Limitations of the Study

This study is recognized to have limitations that could be addressed in future work. The study has been designed, conducted and analyzed by one of the authors; therefore presumptions concerning active and passive physicality will inevitably influence the outcomes. Future work would seek to determine if the notion of active and passive physicality are applicable beyond this study. This is planned in a number of ways; firstly by re-evaluating studies conducted prior to the active and passive physicality notion, secondly by seeking discussion with those involved with interactive prototyping from an academic and commercial context, and finally by evaluating future studies conducted by colleagues.

5 Discussion

In Figures 11 and 12 the IE4 prototype appears to give feedback that is closest to the final iRiver device. These will be discussed along with other, more subtle, differences across the prototypes bringing in comments from the ranking exercise. Observations fall into two categories; recommendations about the design and obstructions caused by the prototype. Recommendations positively help identify how the design can be improved whilst obstructions are caused by features of the prototype that hinder participants in giving meaningful feedback.

5.1 Recommendations about the Design

Physicality of the Dial.
The IE4 prototype was the only prototype that highlighted participants trying to turn the dial to get to the pause function. The physicality of the dial itself could be the cause of this, for the IE4 each rotation has a distinct 'click' which causes a reaction in the interface. However the Arduino prototype did not produce this feedback and its dial had a similar physicality to the final device. This suggests that there must be something else about the prototype that causes the participant to miss feedback for this design recommendation. Several users made comments about the wires of the Arduino prototype being "very distracting" and looking "messier" than the other prototypes, this 'messier' appearance could possibly be the cause of this.

Information Architecture.
The feedback that the interface was longwinded was a common comment from participants of the trial with the final device. The IE4 and Appearance model were both good at drawing the same feedback. The Foam prototype was not able to elucidate this, possibly because the participant was not directly manipulating the prototype and therefore not creating the direct mental link between the physical and digital 'I did not like the fact that I couldn't control the device (interface) from the model'. Meanwhile the Arduino prototype produced few comments about this possibly because the novelty of the prototype itself suppressed the participant's potential frustration with the navigation of the interface "this thing (dial) works alright. I quite like the ability to click". The IE4 seems to give a very direct feel between the interface and interaction, mimicking the final device well. The Appearance model forced the participant to have to continually press the scroll button to navigate the interface, highlighting the sheer number of button presses required to navigate the interface "Very tedious going through all the songs like this".

5.2 Obstructions Caused by the Prototypes

Modeling Physical Interfaces on a Touch Screen.
The Appearance model used a touch screen for the interactive element of the prototype. This prototype gave participants the least difficulties in finding the interactions. Due to the need to represent all the buttons on a touch-screen this prototype clearly indicated where interactions were, even when they were on the side of the device. This made the interactions more obvious for those using this prototype than would otherwise have been. Paradoxically, the very usability of the touchscreen prototype devalues it given the issues users had with the real device.

Obstacles to the Participants Understanding the Prototype.
Figure 11 shows the Foam and Arduino prototypes forced participants to ask for the most help from the facilitator. The Foam model requires the participant to fully engage with the 'speak aloud protocol' because the buttons provide no active feedback. The participant therefore has to wait for the facilitator to operate the interface. In contrast, the Arduino prototype allows the participant to operate it independently, but it may be that the appearance of the wires that seem to the biggest barrier to acceptance.

It may also be that techniques which require the participant to understand the *way* in which the prototype works are not suited for this type of early stage trial.

5.3 Overview of the Four Prototypes

The IE4 Prototype.
The real-time nature and simplicity of this prototype seem to be the important factors in making this prototype the most effective of the prototypes. Participants were able to operate and receive immediate feedback from the interface without an overly complicated looking prototype or altering the scale and form of the model. “I felt very little difference in terms of the final version and white model (IE4) for the quality of interaction - white model (IE4) had a few blips but nothing that is stopping me using the device successfully.” “The addition of working buttons on the prototypes increases the quality of the feel, as the ways in which interaction occurs can be more readily envisioned.”

The Foam Prototype.
This prototype used the ‘speak out loud’ protocol for participants to engage with the interface. Results show that this prototype was less effective at enabling participants to build a mental model of the device resulting in reduced effectiveness of the comments received. “The colour, weight, size and cable connections play a big part of my initial interaction with a product, for this reason the blue foam compared to the final unit was clearly a visual aid as opposed to actual real product comparison.”

The Arduino Prototype.
Participants required more assistance using this prototype. This was a surprise from the most interactive of the prototypes. Participants seemed to be affected by the wires and appearance of this prototype. “The model with blue foam & wires looks messier than the blue foam model but it looks a little bit more functional than the model with blue foam alone.”

The Appearance Prototype.
This prototype used a touch screen to convey the interactions of the prototype. Participants did not identify as many usability errors and had the weakest performance in relation to the final device. This outcome supports Gill et al.’s study in which it was proposed that interactions are easier for a participant to identify on a screen [4]. “Although the silver model (appearance model) looked more like the final version, I did not like the fact that I couldn't control the device from the model and I didn't think having the model alone, without much interaction, was very worthwhile.”

6 Conclusion and Application

The four prototypes trialled in this study explored different aspects of active and passive physicality. The results show that both active and passive physicality are important considerations for early stage user feedback; but it is an even balance of these that produces the most effective prototypes, as seen in the IE4 and Foam prototypes.

Resources should not be used exclusively to ensure the prototype functions well in an electronics and interaction sense (active physicality) if it severely impacts the way the prototype looks or can be held by the user (passive physicality). Likewise, resources spent creating a prototype that looks very close to a final device are not effective if interactions are not well supported.

The IE4 and Foam prototype provided the most accurate data compared to the user experience of the real device. Both the IE4 (£760) and Foam prototype (£60) were of balanced physicality. The Arduino (£1,100) was very strong on active physicality to the detriment of passive physicality whilst the Appearance model (£1,160) was very high on passive physicality but low on active physicality. This suggests that it is those prototypes that are well balanced that are the most effective in this study. Since they are also cheaper they represent strong value for money.[2]

The prototype has long been accepted as a valuable approach to creating valuable and insightful design outputs. However, for interactive devices that have both a physical and digital form, visual fidelity alone is clearly not enough to fully conceive the complete prototype and ensure it will accurately fulfil its purpose. Whilst visual and dimensional fidelity is very much the staple of prototyping, physical fidelity clearly has a role in creating a well-targeted prototype. This study indicates that for interactive prototyping, ‘physicality’ needs to be an even combination of active and passive physicality.

7 Future Work

Future work needs to be conducted to determine if active and passive physicality can be usefully used in assessing prototypes beyond those used in this study. The outcome of this study indicates that a balanced prototype is the most effective. The prototypes used in previous studies [4] [5] should now be assessed in terms of physicality to determine for example if notions of active and passive physicality aid in determining why the data for the ‘flat-face’ prototype differed considerably from the final device. In addition prototypes used in studies by other authors could be categorized to see how they relate to our prototypes.

References

1. Virzi, R.A., Sokolov, J.L., Karis, D.: Usability problem identification using both low and high-fidelity prototypes. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground, Vancouver, British Columbia, Canada (1996)
2. Lim, Y.-K., Stolterman, E., Tenenberg, J.: The Anatomy of Prototypes: Prototypes as Filters, Prototypes as Manifestations of Design Ideas. ACM Transactions on Computer-Human Interaction 15(2) (July 2008)
3. McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B., Vera, A.: Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montreal, Quebec, Canada (2006)

[2] Costs are those shown in Figure 4 minus the software prototyping (shared by all prototypes).

4. Gill, S., Walker, D., Loudon, G., Dix, A., Woolley, A., Ramduny-Ellis, D., Hare, J.: Rapid Development of Tangible Interactive Appliances: Achieving the Fidelity / Time balance. International Journal of Arts and Technology, 309–331 (2008)
5. Hare, J., Gill, S., Loudon, G., Ramduny-Ellis, D., Dix, A.: Physical Fidelity: Exploring the Importance of Physicality on Physical-Digital Conceptual Prototyping. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5726, pp. 217–230. Springer, Heidelberg (2009)
6. Dix, A., Ghazali, M., Gill, S., Hare, J., Ramduny-Ellis, D.: Physigrams: Modelling Devices for Natural Interaction. Formal Aspects of Computing 21(6), 613–641 (2009)
7. Reeves, S.: Physicality, spatial configuration and computational objects. In: First International Workshop on Physicality (2006)
8. Gibson, J.: The theory of affordances. In: Perceiving, Acting, and Knowing, John Wiley & Sons Inc., Hoboken (1977)
9. Norman, D.: The Design of Everyday Things. Basic Books, New York (1988)
10. Norman, D.: The Invisible Computer: Why good products can fail, the Personal Computer is so complex, and information appliances are the answer. MIT Press, Cambridge (1998)
11. Nielsen, J.: Usability Engineering. Academic Press Inc. (1993)
12. Lim, Y., Pangam, A., Periyasami, S., Aneja, S.: Comparative Analysis of High- and Low-fidelity Prototypes for More Valid Usability Evaluations of Mobile Devices. In: Proceedings of NordiCHI 2006 Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles (2006)
13. Maudsley, Greenberg, Mander: Prototyping an intelligent agent through Wizard of Oz. In: Interchi 1993, pp. 277–284 (1993)
14. Gould, J., Lewis, C.: Designing for Usability-Key Principles and what Designers Think. In: CHI 1983, Boston, pp. 50–53 (1983)
15. Gill, S.: PAIPR Prototyping: Some Thoughts on the Role of the Prototype in the 21st Century. In: Prototype: Craft in the Future Tense. Berg (in press, 2013)
16. Burleson, W., Jensen, C.N., Raaschou, T., Frohold, S.: Sprock-it: a physically interactive play system. In: Proceedings of the 6th international Conference on Interaction Design, Aalborg, Denmark (2007)
17. Goodwin, K.: Designing for the Digital Age: How to create human-centered Products and Services. Wiley Publishing Inc. (2009)
18. Dumas, J., Redish, J.: A Practical Guide to Usability Testing (revised edition). Intellect, Exeter (1999)
19. Hertzum, M., Jacobsen, N.: The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods. International Journal of Human-Computer Interaction 15(1), 183–204 (2003)
20. Molich, R., Jefferies, R., Dumas, J.: Making Usability Recommendations Useful and Usable. Journal of Usability Studies, 162-179 (2007)
21. Buchenau, M., Suri, J.: Experience prototyping. In: Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, New York, pp. 424–433 (2000)

Appendix 8: Paperwork for Study Two

Info sheet:

Title of Project: User trials of prototypes of a digital media player

You have been asked to participate in the user trial of a prototype of a digital media player. The purpose of this document is to let you know what this study will involve in order that you may make an informed decision on whether or not you want to take part.

The study will be run by Jo Hare, Ian Culverhouse and Ariana Mihoc, researchers at PDR, a department within the University of Wales Institute Cardiff. The results of the study will be used to inform digital media player design and may also be published in academic papers about design.

By agreeing to take part in this study, you confirm that:

You are over 18 years old

You use a digital media player to listen to music (e.g. iPod, a smartphone, Creative Zen)

There is absolutely no obligation of any kind to join the study and no reasons have to be given for your decision.

What would happen if you join the study?

You will be asked to perform around 5 tasks on a prototype of a digital media player. These tasks will be typical of the normal functions of the music player. The tasks are intended to be a trial of the prototype and not the user.

After the trial you will be asked a series of questions about your experience, there are no right or wrong answers.

We anticipate that the whole study will take between 30 - 40 minutes.

Are there any risks?

We do not think there are any significant risks due to the study. If you feel uncomfortable at any point you are completely free to raise the issue or pull out with no negative repercussions.

Your rights.

Joining the study does not mean you have to give up any legal rights. In the very unlikely event of something going wrong, UWIC fully indemnifies its staff, and participants are covered by its insurance.

Any special precautions needed?

None.

What happens to the results?

Video and audio recordings of the research will be studied and transcribed. We will then look for reoccurring themes, values and views.

What happens if you want to change your mind?

If you decide to join the study you can change your mind and stop at any time. We will completely respect your decision. There are absolutely no penalties for stopping.

How we protect your privacy:

Your name and any other personal details will be kept separately from any other documented research and we will take steps to ensure that no one can identify you from the research findings.

Once we have finished the study we will destroy all of the audio and visual recordings of the studies that may identify you. And any transcription will be coded and kept completely separately from your personal details.

We keep a copy of your name along with your consent form for 5 years as we are required to do so by UWIC.

PLEASE NOTE: *YOU WILL BE GIVEN A COPY OF THIS SHEET TO KEEP, TOGETHER WITH A COPY OF YOUR CONSENT FORM*

Active and passive physicality: making the most of low fidelity physical interactive prototypes

Joanna Hare*

The National Centre for Product Design and
Development Research (PDR),
Cardiff Metropolitan University,
Cardiff, CF5 2YB, UK
E-mail: jhare@pdronline.co.uk
*Corresponding author

Steve Gill and Gareth Loudon

Cardiff School of Art and Design (CSAD),
Cardiff Metropolitan University,
Cardiff, CF5 2YB, UK
E-mail: sjgill@cardiffmet.ac.uk
E-mail: gloudon@cardiffmet.ac.uk

Alan Lewis

The University of Wales,
King Edward VII Avenue,
Cardiff, CF10 3NS, UK
E-mail: ProfAlan.Lewis@wales.ac.uk

Abstract: This paper presents three case-studies which comprise a systematic investigation into the use of low fidelity physical interactive prototyping techniques to form a design principle based on the constructs of active and passive physicality. It proposes that, with a better understanding of active and passive physicality, designers can make more effective prototypes for early stage user trials. Results of our studies indicate that the most effective prototypes balance both active and passive physicality equally. In addition, the notion of physicality can demonstrate why, in our studies; paper prototyping, screen-based prototypes and even Arduino prototypes produced unsatisfactory user data.

Keywords: physicality; fidelity; prototypes; interactive prototypes; computer embedded devices; user testing; usability; human computer interaction; industrial design; paper prototyping; design research.

Reference to this paper should be made as follows: Hare, J., Gill, S., Loudon, G. and Lewis, A. (xxxx) 'Active and passive physicality: making the most of low fidelity physical interactive prototypes', *J. Design Research*, Vol. X, No. Y, pp.000–000.

Biographical notes: Joanna Hare is a Researcher and Product Designer at the National Centre for Product Design and Development Research. She is part of the User Centric Design Group which undertakes academic research and commercial interaction design projects. Her specific interest is developing tools and techniques for user research that can be applied in a commercial setting. She has been an RA on the AHRC/EPSRC funded 'Designing for Physicality' project investigating the impact of physicality on product design – how humans experience, manipulate, react and reason about 'real' physical things, and how this understanding can inform the future design of innovative products.

Steve Gill is Professor of Interactive Product Design and Associate Dean (Research) in Cardiff School of Art & Design. He has 20 years' experience in industry and academia and has published widely in journals, book chapters and conference proceedings. His research group has a strong record in applied research, enterprise and consultancy. He has been active in the design research community for many years, collaborating with high profile academics, notably Prof. Alan Dix at the University of Birmingham with whom he is co-authoring a book, *Touch IT*, on the importance of physicality in the design of computer embedded products.

Gareth Loudon is a Principal Lecturer in Product Design at the Cardiff School of Art and Design and co-Founder and Director of the Centre for Creativity Ltd. that undertakes research, training and consultancy in key areas of creativity. His research interests focus on creativity and the product innovation process, combining ideas from anthropology and psychology, engineering and design. He has over 25 years' experience in academic and industrial research and has taken several research ideas all the way through to commercial products for large companies such as Apple Computer. He has several patents to his name and over 40 publications in total. He is a Chartered Engineer, a Fellow of the Institution of Engineering and Technology and a Fellow of the Higher Education Academy.

Alan Lewis was formerly the Director the National Centre for Product Design and Development Research (PDR) and the Dean of Research at the Cardiff Metropolitan University (formerly UWIC). He has worked with many companies in a variety of industry sectors helping them to improve their product development processes. He has a particular interest in the management of product design and development processes within the SME sector of the economy and has published extensively in this field of research.

This paper is a revised and expanded version of a paper entitled 'Physical fidelity: exploring the importance of physicality on physical-digital conceptual prototyping' presented at Interact 2009, Uppsala, Sweden, 24–28 August 2009; and a paper entitled 'The effect of physicality on low fidelity interactive prototyping for design practice' presented at Interact 2013, Cape Town, S. Africa, 2–6 September 2013.

1 Introduction

This research brings together knowledge from a variety of disciplines to apply them specifically to the construction of physical interactive prototypes of computer embedded devices. Computers have been embedded in our products for nearly half a century, and with an ever evolving stream of new technology, computer embedded devices are a fast-paced topic but this research draws on the essence of the designed object and rises above the latest technology in order to explore physicality at a fundamental level.

The inclusion of an 'invisible' computer in computer embedded devices requires electronic interpretation of our interactions for any output. This electronic interpretation can 'violate' the basic principles of physical objects (Ghazali and Dix, 2005) and does not have to conform to our experience of the physical world. Computer embedded devices represent an interesting challenge for designers; typically the physical product is designed by an industrial (or product) designer while the software interface is the domain of the human-computer interaction (HCI) specialist. Both these disciplines, although obviously interdependent, are historically very different resulting in a lack of coherence in our computer embedded devices (Overbeeke, 2013).

An iterative design process is considered an effective design process (Rubin and Chisnell, 2008). This approach advocates rapid iterative user testing through inter alia usability trials (Nielsen, 1993), semi-structured interviews (Sharp et al., 2007), and expert reviews (Molich and Jeffries, 2003). Low fidelity prototypes are a fundamental tool for many for these techniques, and interactive prototypes can be used to explore the digital considerations within the physical form of computer embedded devices.

It is these low fidelity interactive prototypes, fundamental to the iterative design process, which this research focuses on. By ensuring these prototypes are effective during studies the design team can make better design decisions and thus more usable products.

2 Fidelity

When creating a prototype, the designer needs to balance the visual and functional needs of the prototype, the environment within which it needs to operate (for example, user trials, demonstration, or talk through), and the skills and resources of the prototype team (time and equipment). This balance will have significant impact on the fidelity of the prototype.

Virzi (1989) describe fidelity as being "a measure of how authentic or realistic a prototype appears to the user when it is compared to the actual service". Rudd et al. (1996) characterise low fidelity prototypes as "limited function, limited interaction prototyping efforts *…+ constructed for illustrating concepts, design alternatives and screen layouts". The authors continue by defining high fidelity prototypes as being 'fully interactive' meaning that a user can "interact with the user interface as though it is a real product". Nilsson and Siponen (2006) propose that fidelity can be defined by the response of the prototype, from fully automatic (user-driven) to non-automatic (facilitator driven). McCurdy et al. (2006) proposed five 'dimensions' of fidelity that can be defined as somewhere between high and low within the same prototype, namely, aesthetics, depth of functionality, breadth of functionality, richness of data and richness of interactivity.

When designing computer embedded devices, physical prototypes are required that incorporate the digital interaction. Prototypes have traditionally been referred to by their fidelity, yet research into fidelity has been predominantly on software only prototypes (McCurdy et al., 2006; Nielsen, 1993; Rudd et al., 1996. Of the research that does focus on physical interactive prototypes, the construction of the physical prototype is rarely typical of the product design process. For example, Lim et al. (2008) use a real mobile phone and vary the level of fidelity of the on-screen interaction; Virzi et al. (1996) use a paper keyboard but not a physical model for their electronic book and Sauer et al. (2010) overlaid a cardboard mock-up over the real appliance.

Our research set out to determine if a better understanding of physicality could provide the means of creating prototypes which are effective at eliciting meaningful comments and insights during early stage user trials. We define ‘meaningful’ comments as those that focus on improving the overall intended design of the concept as opposed to the interface in isolation or the construction technique of the prototype.

3 Physicality

Physicality is central to our experience of computer embedded devices, from how we exist in our bodies within the physical world, to how we perceive interactions with the physical world, and the point at which we interact with that physical world. Literature points towards three philosophical discussion areas related to physicality and the designed object; humans as physical beings within our physical world (embodiment and phenomenology) (Clarke, 1998; Dourish, 2001; Merleau-Ponty, 1945; Haugeland, 1998), perception of interaction through physical signifiers (affordances) (Gibson, 1979; Norman, 1998; McGrenere and Ho, 2000; Gaver, 1991) and at the point at which the digital and physical meet (interaction) (Ghazali and Dix, 2005; Dourish, 2001). Thus, we define physicality as the physical aspects or qualities of both an object and its interaction; this includes our physical bodies in relation to that object.

4 The constructs of active and passive physicality

We propose that the physicality of a prototype can be considered on two levels; that of active and passive physicality where; passive physicality is the perceived affordance based on the visual appearance and tangibility of the prototype, and active physicality is the perceptible experience of interacting with the prototype.

To explain these terms a useful starting point is that of Dix et al. (2009), who regard a physical device removed from its context, and ‘separated’ from its digital operation, in order to consider the mapping of the device ‘unplugged’. This is the basis of ‘passive’ physicality; the judgements that can be made about a device by considering both its visual appearance and its tangibility (by touching it), without switching it on.

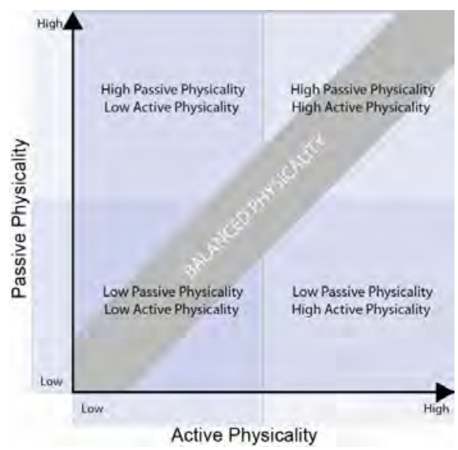
Assumptions are formed about the physicality of the device-based purely on its visual appearance as Reeves (2006) demonstrates by asking; do you grasp a cup by its handle or by the body? Decisions are made about the comfort of the cup’s handle by its appearance and the perceived weight of the contents of the cup. The tangible nature of the prototype is also a key aspect of passive physicality; this includes the way the device feels in your hand, its weight, the location of any interactions and surface finish.

Passive physicality also has its roots in Norman’s definition of affordances (Norman, 1998). Affordances suggest ways of interaction which are dependent on the user’s ability to perceive it. The intended design of the device has affordances; in addition, the way in which the prototype is constructed brings its own, different, affordances that affect the way in which the user perceives the prototype. If the prototyping technique used interferes with the user’s ability to make a mental model of how the prototype is operated this will impact passive physicality. For example, if interactions are ‘hidden’ by the physical prototyping method, users cannot perceive that an interaction is possible. Passive physicality forces the designer to recognise that the way in which the physical prototype is executed has an impact on the user's experience of that prototype.

Active physicality is concerned with the physical act of interacting with a prototype in its 'on' state which thus requires electronic interpretation of the action resulting in feedback that can be perceived. Tangible feedback comes from the 'feel' of the interaction, whether it is the simple 'bounce-back' of a button or electronically enhanced haptic feedback. This interaction will cause the electronic state of the device to change perceptibly, for example a screen change, a light coming on or a mechanism engaging.

The boundary between active and passive physicality is the point at which manipulation of the device occurs which requires electronic interpretation or mechanical action (or both). For example, we use our sense of touch to determine whether buttons fall in a 'natural' location (passive physicality) but if we then interact with those buttons to determine what they do and how they feel, this now falls under active physicality. If those actions are intended to initiate further actions, for example changing a screen element, this should be considered alongside its tactile feedback. An interaction which does not comprise all of its intended actions will have lower active physicality than one that does. For example, if a switch is not connected it will deform and feel like it should but it will not result in any feedback beyond the tactile.

Figure 1 Graph showing passive versus active physicality levels (see online version for colours)



We propose that both active and passive physicality can be considered on a scale of low to high (Figure 1). Our case studies suggest that prototypes which fall below certain levels of either active or passive physicality in relation to the design intent are least effective, and prototypes that balance active and passive physicality equally are the most effective. In this situation an 'effective' prototype is one which elicits feedback related to the intended design to enable the next iteration of the design to take place.

The proposal that active and passive physicality should be 'balanced' recognises that many prototyping construction techniques require a compromise of some kind. For example, the use of electronics within a prototype necessitates components and power requirements which could impact the size of the prototype and the demand for a highly realistic prototype could impact the way in which the prototype can be interacted with. In these scenarios the resultant physicality of the prototype is affected even though its fidelity is not necessarily altered, without an understanding of this affect any prototype created could be limited in its effectiveness.

The design of these case studies was based on two independent variables; the design of the device and the structure of the trials. In each study the prototypes were constructed of the same design intent with the same functions and features and each prototype within the study was trialled in an identical manner. In addition, each of the case studies had a specific independent variable which determined prototype construction, these were; physicality levels for the media player (case study 1), decreasing fidelity levels for the home phone (case study 2) and time limitations for the photo management device (case study 3). The resultant prototypes were dependent on these parameters and the impact on physicality could then be assessed. Once physicality levels have been determined and the prototypes have been trialled with a consistent structure, any differences in the results of the user trials can be compared in relation to the physicality of the prototype.

5 Case study 1 – media player

In this study, the technique used to construct the four prototypes was determined by our proposed definition of active and passive physicality (Hare et al., 2013). The intention was to include the four permutations of active and passive physicality levels as demonstrated by each quadrant of Figure 1. Subsequently, this case study provides an illustrative example of active and passive physicality.

5.1 The prototypes

An existing media player was chosen as the basis for the construction of four prototypes. A single interface was coded in Adobe Flash for all prototypes and adapted to the needs of each. Preparatory work ensured that this interface would be suitable for all prototypes and that the adaptation of the interface was possible for all. The prototypes are shown in Figure 2.

The ‘blue foam’ prototype was constructed from model-making foam. Interaction was based on the Wizard of Oz technique (Maulsby et al., 1993), the interface was operated remotely by the facilitator, the participant was asked to follow the ‘think out loud’ protocol (Gould and Lewis, 1983), the facilitator could react to what the participant was saying and interacting with on the foam prototype. *Construction time: six days 2 hours*.

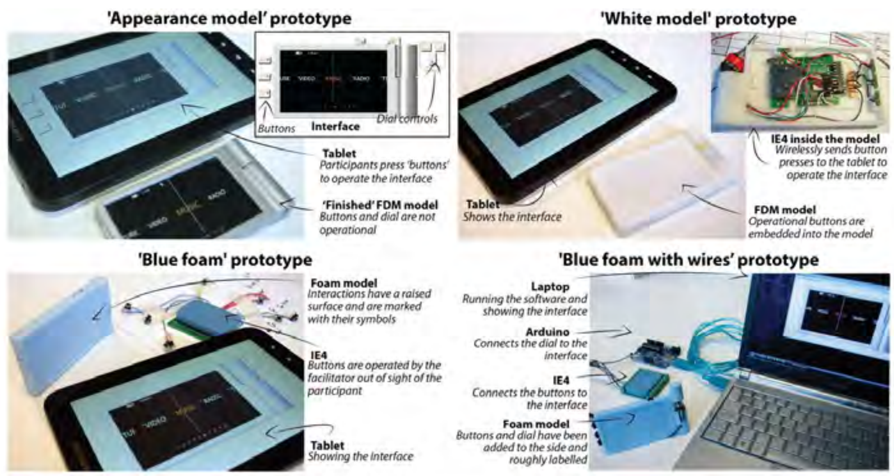
The ‘white model’ was constructed using rapid prototyping techniques; it was very similar in size and shape to the final device. The buttons and the dial were integrated to make the prototype interactive. An IE4 (Gill, 2013) was used to connect the buttons to a laptop. The Flash interface, shown on a tablet, ‘listens’ for key presses from the IE4 and triggers changes in the interface when the participant interacts with the prototype. *Construction time: ten days 2 hours*.

The ‘appearance model’ was intended to reflect the final appearance of the device as accurately as possible. The form was constructed using rapid prototyping techniques and finished to facsimile level. The Flash interface was operated by the participant on a separate touch screen tablet. *Construction time: ten days*.

An approximate foam model was constructed for the ‘foam model with wires’ to accommodate the off-the-shelf buttons and dial. The dial was connected to an Arduino

(Burleson et al., 2007) which received the analogue signals and sent them to the computer running the Flash interface. The buttons were connected to an IE4. Due to the extra code required for the Arduino, the interface was shown on a laptop. *Construction time: eight days 6 hours*.

Figure 2 Overview of the media player prototypes (case study 1) (see online version for colours)



5.2 Assessing physicality

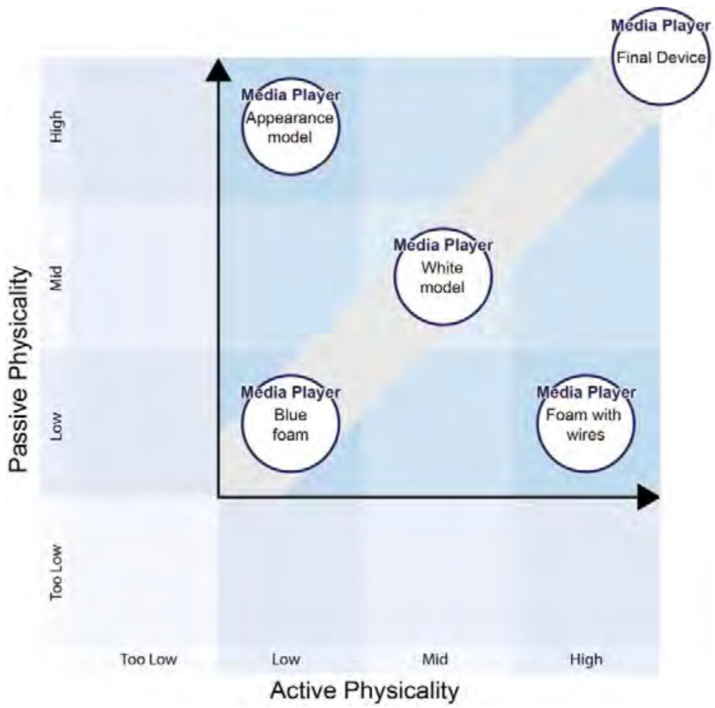
The levels of physicality are shown in Figure 3. The tangible and visual qualities of the blue foam prototype are accurate but low fidelity (low passive physicality). Interaction is based on the 'speak out loud' protocol and operated by the facilitator, buttons are cardboard but the dial does rotate (low active physicality).

The white model was similar to the final device in its form, weight and the location of buttons, therefore this prototype has higher passive physicality than the blue foam prototype. Upon interaction, the haptic and visual feedback is good approximation of the final device, resulting in higher levels of active physicality in relation to the blue foam prototype.

The appearance model was an accurate representation of the final device (high passive physicality). However, the absence of electronics means there was no feedback from the buttons or dial; and the interface was operated on a touch screen separate to the physical prototype resulting in low active physicality. This prototype really brings out the distinction between active and passive physicality because the buttons on the model have good haptic feedback, yet this does not raise the assigned active physicality level significantly because they do not function.

The foam model with wires has an approximate physical model that is clearly modified to accommodate the switches and dial; the wires are very apparent and visually impact the prototype resulting in low passive physicality. Upon interaction, the feedback of the interactions accurately represents the final device (high active physicality).

Figure 3 Assessment of physicality for the media player prototypes (case study 1) (see online version for colours)



5.3 Results of the user trial

Users were asked to perform five tasks before commenting on the main menu options, this ensured each participant had the same knowledge of the device for a semi-structured interview. The data was analysed to elicit design recommendations for each prototype and these were compared to the final device (Hare et al., 2013).

Participants using the white model gave good feedback indicative of the final device. Results of the blue foam prototype show that this prototype was less effective at enabling participants to build a mental model of the device resulting in reduced effectiveness of the comments received. Participants struggled to relate the action they were performing on the physical model to what was happening onscreen (active physicality).

Participants using the foam model with wires required more assistance using the prototype. This was a surprise given that this prototype had the highest active physicality levels. Participants seemed to be affected by the wires and appearance of this prototype (its passive physicality) resulting in less meaningful comments. The appearance prototype had the weakest performance; although some interesting comments were received, the comments elicited by this device did not accurately reflect those of the final device.

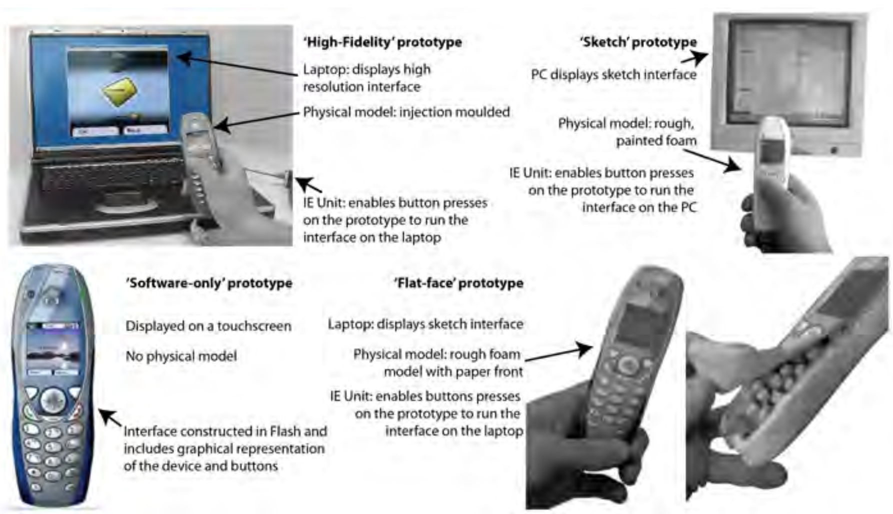
6 Case study 2 – mobile home phone

In this study, the technique used to construct the prototypes was determined by fidelity levels. The study set out to discover if results of a user trial with a tangible prototype were more similar to the final product than a software-only prototype, and the subsequent level of fidelity required of this prototype (Gill et al., 2008). In total, four prototypes were constructed and compared to a final device. This case study demonstrates how an understanding of active and passive physicality has provided a framework by which to better understand unforeseen results.

6.1 The prototypes

Four prototypes were constructed; the first two being a high fidelity model and a software-only prototype (mimicking common prototyping practices), and a further two that lowered the level of fidelity of the prototype, these are shown in (Figure 4).

Figure 4 Overview of the home phone prototypes (case study 2) (see online version for colours)

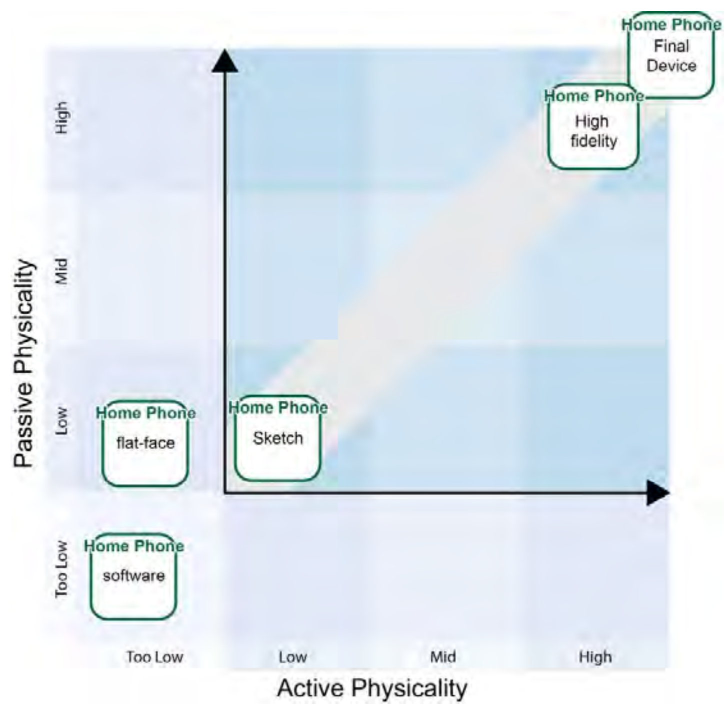


The 'high fidelity' prototype was created by connecting an IE unit to buttons in the casing of the final device, the IE unit enabled button presses on the phone to trigger a mock-up of the phone's interface created in Flash and shown on a laptop. The same Flash interface was used for the 'software-only' prototype and operated through a touchscreen laptop. The 'sketch' prototype consisted of a blue foam model with basic integrated buttons and sketch graphics within the Flash interface. The IE unit was again used to connect the physical model to the computer to operate the interface. A blue foam model was created for the 'flat-face' prototype, instead of embedding the buttons into the front of the phone (one of the more time consuming tasks involved in creating the prototype) a paper print out covered the physical buttons; the same sketch Flash interface was used.

6.2 Assessing physicality

The physicality levels are shown in Figure 5. The tangible and visual qualities of the physical model of the 'high fidelity' prototype are very similar to the final device, with the difference in weight and appearance of the wires (connecting to the IE unit) being the only compromises (high passive physicality). Upon interacting with the prototype, the buttons have the same feel as the final device with the onscreen graphics performing to a high fidelity albeit on a remote screen (high active physicality).

Figure 5 Assessment of physicality for the home phone prototypes (case study 2) (see online version for colours)



The visual appearance of the 'sketch' prototype is very crude but the tangible aspects of scale, form and button location are a good approximation of the final design (low passive physicality). Upon interaction, the buttons have the similar feel as the final device, the onscreen graphics were very crude in appearance but the structure of the interface is identical to the high fidelity prototype (low active physicality).

The scale and form of the 'flat-face' prototype are restricted due to the front being removed, the printed visual appearance is reasonable and the buttons appear to be in a good approximate location (passive physicality is marginally lower than the 'sketch' prototype). Yet upon interaction it becomes apparent that the 'hit' area of the buttons differs from what is visible on the surface, in addition the physical feedback of the buttons was reduced by the paper, the interface was identical to the sketch prototype (active physicality is significantly lower than the 'sketch' prototype).

There was no tangible model for the 'software-only' prototype; therefore the only concession to passive physicality is a two-dimensional graphical presentation of the design resulting in extremely low passive physicality levels. The interface was identical to the 'high fidelity' prototype yet interaction with the interface was vastly different to the final device with no tactile feedback of the device or buttons. This marks a very interesting attribute of active physicality, the lack of a physical device to hold and manipulate has a very marked effect on active physicality levels despite the onscreen interface being considered 'high fidelity'.

6.3 Results of the user trial

User trials were conducted utilising the four prototypes and the final device. Users were asked to complete six tasks and the success rate of each task was recorded (Molich and Dumas, 2008) along with the time taken to complete the task. The 'high fidelity' prototype produced similar results to the final device, significantly outperforming the 'software-only' prototype. The 'sketch' prototype was found to perform similarly to the final device. The performance of the 'flat-face' prototype however, was significantly reduced. It appeared that the flat face of the prototype did not replicate the true physicality of the product sufficiently, and the result was more user error which produced in slower performance times and worse performance ratings.

The initial publication of this study concluded that it is not the level of fidelity that is important but rather the considerations of tangibility and physicality. It proposed that there was something which was lacking in the physicality of the 'flat-face' prototype that prevented it from being an effective prototype. When our hypothesis of active and passive physicality is considered, it becomes apparent that the 'software-only' prototype has no passive physicality and little active physicality. What is surprising in this case is that despite the interface being identical to the high fidelity prototype there is a difference in way in which the interface is operated. There is no physical model or buttons with which to operate the interface, and this has a significant impact on the active physicality level because the user cannot tangibly feel the model or interaction.

In this trial, the 'sketch' prototype, although low fidelity, implements enough active and passive physicality for the user to understand the design on a similar level to the 'high fidelity' prototype. This reveals a significant saving in time and expense in terms of constructing a prototype, in addition to being able to construct this type of prototype earlier in the design process enabling more iterations of the design.

On initial appraisal, the 'flat-face' prototype appeared as though it would produce effective results because the only difference between this and the sketch prototype is the paper covering the buttons. But when notions of active and passive physicality are applied, it becomes apparent that active physicality is very low in comparison to the design intent. Feedback of interaction is poor since the participant cannot determine exactly where the 'hit area' is underneath the paper, resulting in unsatisfactory feedback upon interaction (active physicality). In addition, it seems that the interactions are not transparent enough for the user to understand how to operate the prototype, in other words, the appearance and tangibility of the prototype suggest there is little the participant can do with the prototype (perceived affordance resulting from passive physicality). This prototype has been a really interesting case study because it seems to marginally challenge the boundaries of an acceptable low fidelity interactive prototype.

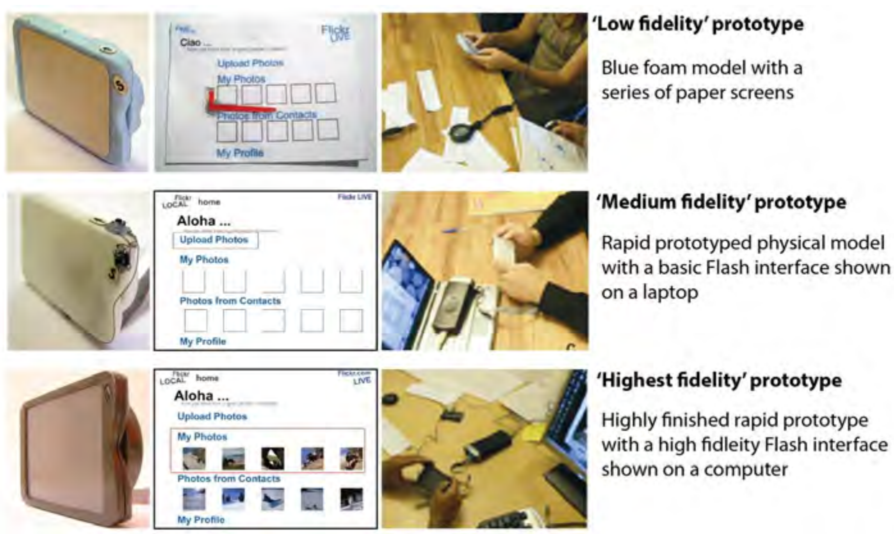
7 Case study 3 – conceptual photo management device

In this study, the technique used to construct three prototypes was determined by allocating time limits during construction. A conceptual device was chosen in order to fully reflect the nature of low fidelity prototyping early in the design process when there are many unresolved aspects of the design. Initial design work was undertaken in order to reach a stage where, in a real design process, an interactive prototype would be the next natural step. Each of the resulting prototypes used this initial design work as the starting point, therefore only the time to construct the prototype differed (Hare et al., 2009). This case study further demonstrates how the framework of active and passive physicality can be used to better understand unforeseen results.

7.1 The prototypes

Three prototypes were constructed (Figure 6); the considerations that drove the level of fidelity and its effects on the physicality were purely time-based, with the allocated times of 4 hours, 14 hours and five days.

Figure 6 Overview of the conceptual photo management device prototypes (case study 3) (see online version for colours)



A blue foam model was created for the 'low fidelity' prototype. A series of paper screens were created for the interactions based on the principles of paper prototyping. The prototype required a facilitator to operate the 'interface' while the user talked through their interactions. *Construction time: 4 hours*.

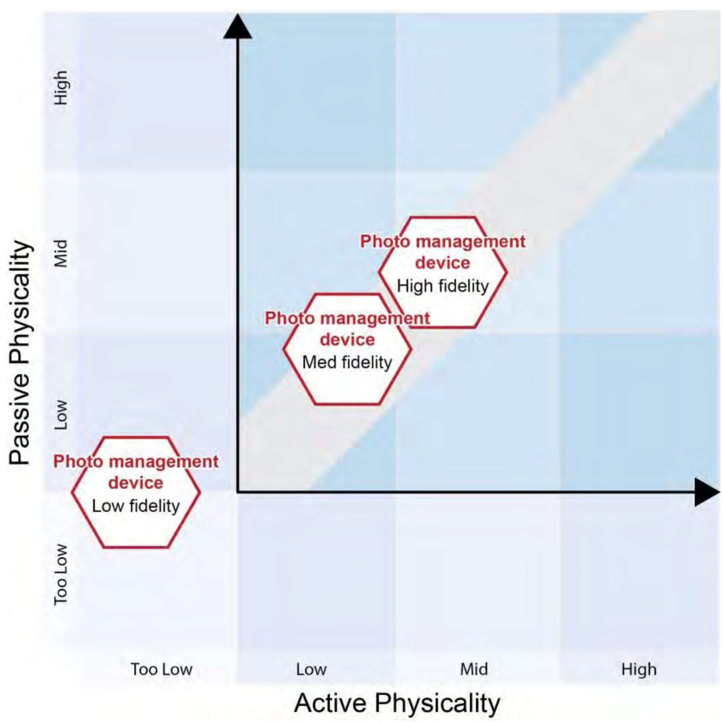
The physical model was created using rapid prototyping techniques for the 'medium fidelity'' prototype. Buttons were integrated into the physical model and connected to an IE Unit operating a Flash interface displayed on a laptop. *Construction time: 14 hours*.

The 'highest fidelity' prototype was based on the mid-level with the extra time used to give the physical prototype a more realistic visual finish and improve the feel of the interaction. The interface was developed to operate in a smoother manner reflecting the intended design. *Construction time: five days*.

7.2 Assessing physicality

The levels of physicality are shown in Figure 7. Although the physical form is relatively accurate for the low fidelity prototype, it feels very lightweight; interactions are clearly depicted but perceptibly non-functional (low passive physicality). Interaction relies on the participant pressing cardboard buttons and talking through their actions with the facilitator interpreting this by adjusting the paper screens. Although buttons are accurately located on the prototype, there is little tactile feedback of the buttons and delayed visual feedback of the interface (very low active physicality).

Figure 7 Assessment of physicality for the photo management concept prototypes (case study 3) (see online version for colours)



The physical form factor of the medium fidelity prototype is relatively accurate; the unfinished form and tacked on buttons inform the user that interaction is possible but it does not visually reflect the final device. Therefore the passive physicality of this prototype is low but still higher than the low fidelity model. The dial gives haptic feedback but this is not representative of the intended design; this dial feels 'clunky' and cannot rotate 360 degrees whereas the intended design fully rotates giving more subtle haptic feedback. Visual feedback of the interaction is immediate and the interface is functionally accurate but screen animations are not as refined as the intended design, therefore active physicality is higher than the low fidelity model. This example demonstrates the importance of relating the physicality of the prototype to the *intended* design of the device. In this case the active physicality of the prototype would have been higher if the intended design had reflected the haptic qualities of the dial mechanism used in the prototype. Yet a more representative dial was known to significantly impact the length of time this prototype would have taken to construct because of the extra coding required as demonstrated by the highest fidelity prototype.

The highest fidelity prototype was constructed and finished to accurately represent the intended design visually and tangibly (high passive physicality). It could be further improved by ensuring the weight of the device is more accurate. The interactions of the device reflect the intended design well with the dial providing full rotation with subtle haptic feedback and the interface includes good visual feedback (high active physicality).

7.3 Results of the user trial

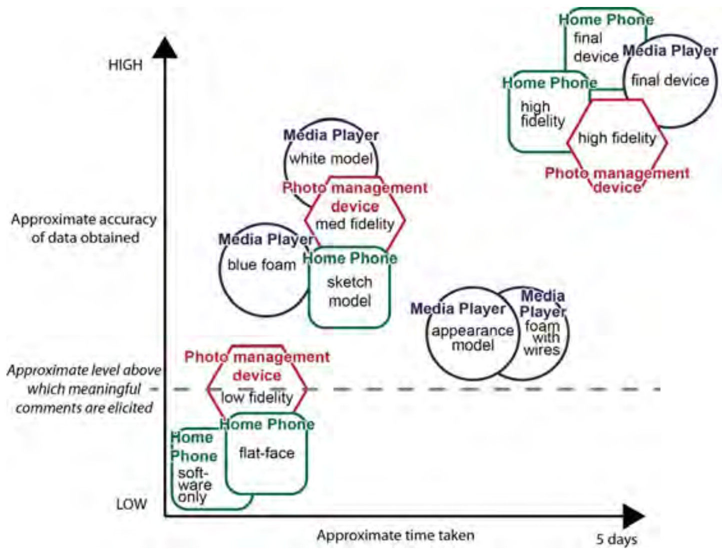
Users were asked to perform five tasks on the prototypes, task success rate was recorded and discourse analysis was performed on the resulting data. Initial analysis showed that task success rate did not differ significantly across the prototypes, although results suggested that the greatest difficulty for users of the lowest fidelity prototype was identifying the correct interaction; whilst users of the medium and highest fidelity prototypes had more problems creating a 'mental model' of the interface. Discourse analysis revealed that the 'medium' and 'high' level prototypes were more effective at eliciting useful user comments than the 'low fidelity' prototype.

When our hypothesis of active and passive physicality is considered, we can see that, despite being of very different fidelity, the physicality of the medium and high fidelity prototypes is relatively similar. Therefore, in terms of physicality, very little has been added to the high fidelity prototype despite the additional time spent creating the prototype. The low fidelity prototype, however, is very low in active physicality due to the lack of haptic and visual feedback of the prototype. Despite setting out to assess the effect of physicality, when the notions of active and passive physicality are applied the prototypes used in this study are fairly similar. This could explain why the results of this study were inconclusive.

8 Discussion

In total, 11 prototypes of three separate products were studied. The relative success of each prototype can be approximately determined by comparing the data the prototype produced during user trials to the other prototypes in that case study. Figure 8 shows the relative success of the prototype versus the time taken to create the prototype (for study two these are approximate times).

Figure 8 Relative success of the prototype versus the time taken to create the prototype (see online version for colours)



8.1 Prototypes without embedded electronics

The 'software-only' prototype of study two took the least time to create. This prototype had no physicality, although it could be argued that there is some physicality if the intended device is touchscreen and it is prototyped on a touchscreen. However, this was not the case in our study and the inclusion of an 'appearance' model can be used to address this. Our 'appearance' prototype (study three) was close to the design intent for passive physicality but active physicality remained low. Previously, it was thought that the inclusion of an 'appearance' model was adequate to inform the design process, but our studies have demonstrated that these prototypes produced unreliable data compared to the other prototypes in the series. In study one, the interface required between 59% and 96% of prototyping time depending on its level of fidelity. Therefore, with as little as 4% extra time, the effectiveness of the prototype can be greatly improved by the inclusion of an interactive physical model bringing the level of active versus passive physicality into balance.

The effectiveness (or lack thereof) of paper prototyping has been a surprise outcome of these studies. It is a technique used regularly in commercial work, yet study one suggests that the lack of real-time feedback (active physicality) results in a decrease in the quality of results. This prototyping technique is classified as a 'non-automatic' by Nilsson and Siponen (2006) because the facilitator plays a very noticeable role in the eyes of the participant. The reduction in the quality of data was thought to be due to the delay of the facilitator in updating the interface and the inability to 'explore' the interface because of the added 'unnecessary' work the participant felt they were causing the facilitator (Sefelin et al., 2003). Indeed, Nielsen (1990) found a similar result where users found significantly less 'global' problems when using a paper prototype compared to a software prototype. Yet, paper prototyping has been proven to be a successful method for usability studies (Snyder, 2003; Sefelin et al., 2003). The user trials of study three were designed to obtain feedback about the scope of the overall design (what Nielsen describes as ‘global’ considerations) rather than task structure. Paper prototyping seems to be more appropriate when exploring the more detailed information architecture of an interface. Perhaps this suggests a lower limit of active physicality; the prototype should appear to be ‘automatic’, or real-time, for early stage user feedback based on usability trials.

Study one addressed the lack of active physicality in the paper prototype through the ‘blue foam’ prototype; this increased levels of active physicality through the facilitator operating an ‘automatic’ interface thus balancing the levels of active and passive physicality. This ‘blue foam’ prototype proved successful in eliciting reliable user feedback; the higher fidelity ‘white model’ outperformed this foam prototype but as a quick and dirty prototype this ‘blue foam’ prototype was a success.

The ‘appearance’ prototype of study one posed an interesting question in relation to the buttons. On this prototype the buttons felt similar to the end device but they were not functional. Active physicality has been proposed to be the perceptible feedback of interacting with the device; and, haptically at least, the interactions are accurate. Yet these interactions do not trigger any other feedback, so the user is not able to relate their interactions to the product as a whole.

8.2 ‘Smart’ prototypes

The inclusion of electronics within the prototype is a common way to increase the fidelity of interactive prototypes; this enables real-time interaction and an improved richness of interactivity (impacting active physicality). Seven of our prototypes covered a variety of approaches to making the prototype ‘smarter’. Some of those approaches have resulted in an adjustment to the physical form and some have resulted in additional wires being present; in all of the prototypes studied, the screen was outside the physical model. Studies one and two demonstrate that the remote screen had no impact on the data gathered by comparing results to the final device with integrated screen. This allows a significant reduction of the development time of prototypes pushing levels of fidelity and physicality even lower. The prototypes that had a significant impact on passive physicality produced the least reliable data, the two extreme cases in our studies were the ‘flat-face’ prototype of study two and the ‘foam model with wires’ of study one. The physical form of the ‘foam model with wires’ was distorted due to the size of the switches and dial used, in addition, the wires and prototyping board were clearly visible impacting passive physicality. Participants commented that they felt ‘intimidated’ by the appearance of the electronics and that interactions were not easy to reach on the prototype. The ‘flat-face’ prototype used in study two had a paper cut-out covering the buttons to avoid the need to embed the buttons in the front of the model, saving a few hours’ work. The effect of this paper cut-out was two-fold; firstly the level of passive physicality was too low because the ‘hit-area’ shown on the paper cut out was not the true hit area of the buttons beneath it, and secondly, upon interacting with the device (active physicality) the paper and misalignment of hit-areas caused inadequate feedback.

8.3 Balancing physicality

In our studies, the most successful prototypes balanced both active and passive physicality equally, these included the very low fidelity ‘sketch’ prototype and ‘high fidelity' prototype of study two, the low fidelity 'blue foam' prototype and the higher fidelity 'white model' of study one. The least successful prototypes did not address one, or both, aspects of physicality, these were the 'software-only' and 'flat-face' prototypes of study two, and the 'low fidelity' paper prototype of study three. The remaining prototypes focused too much on either active or passive physicality with less consideration of the other, although these prototypes produced valid data, it was not as reliable as the well-balanced prototypes.

When we relate these notions back to the time taken to create each prototype we can see that the extra time invested in the prototype was perhaps inefficient. For example, the 'foam model with wires' and 'appearance' models of study one took more time than the 'white model' of that same study but were less successful. In order to increase the effectiveness of the 'appearance' prototype and 'foam model with wires', they could be combined with further investment to source buttons and dials that did not have a significant impact on passive physicality. This type of investment would be more justifiable towards the later stages of the design process.

As was hypothesised in study two, it seems that it is not the level of fidelity that is important in these prototypes. Rather it is considerations about the physicality of the prototype in relation to the design intent, specifically that there is a good balance between active and passive elements of the prototype.

9 Conclusions

We propose that the physicality of the prototype should be considered on two levels; that of active and passive physicality where passive physicality is the perceived affordances based on the visual appearance and tangibility of the prototype and active physicality is the perceptible experience of interacting with the prototype. This notion of active and passive physicality has provided a clearer understanding of the results obtained in our investigations.

Physical interactive prototypes require an electronic prototyping platform, software, interactions (such as button and sliders) and hardware (to run the prototype) within a physical form. Many different prototyping techniques exist that bring together these elements in a variety of ways; the application of active and passive physicality in the planning stage enables prototypes to be executed in the most efficient manner to elicit meaningful comments and insights from user trials. Some of the prototypes presented in this paper push physicality to a level where results were compromised, suggesting that there is a certain level of physicality that prototypes should not fall below. The most successful balanced the levels of active and passive physicality equally. Therefore, resources should not be used exclusively on the prototypes interaction (active physicality) if it severely impacts the ways the prototype looks or can be held by the user (passive physicality). Likewise, resources spent creating a prototype that closely resembles a final device is not effective if interactions are not well supported.

10 Future work

Future work will seek to determine the relevance of passive and active physicality beyond our case studies by evaluating prototypes emerging from both research and commercial projects. Further case studies could focus on different prototyping techniques such as augmented and virtual prototyping plus devices that change shape such as those with flexible screens.

References

Burleson, W., Jensen, C., Raaschou, T. and Frohold, S. (2007) 'Sprock-it: a physically interactive play system', *Proceedings of the 6th International Conference on Interaction Design*, ACM, Aalborg, Denmark.

Clarke, A. (1998) 'Where brain, body and world collide', *Daedalus: Journal of the American Academy of Arts and Science*, (Special Issue on the Brain), Vol. 127, No. 2, pp.257–280.

Dix, A., Ghazali, M., Gill, S., Hare, J. and Ramduny-Ellis, D. (2009) 'Physigrams: modelling devices for natural interaction', *Formal Aspects of Computing*, Vol. 21, No. 6, pp.613–641.

Dourish, P. (2001) *Where the Action Is: The Foundations of Embodied Interaction*, MIT Press, Cambridge, Massachusetts.

Gaver, W. (1991) 'Technology affordances', *SIGCHI Conference on Human Factors in Computing Systems Reaching Through Technology*, pp.79–84.

Ghazali, M. and Dix, A. (2005) 'Knowledge of today for the design of tomorrow', *Proceedings of the 2nd International Design and Engagibility Conference (IDEC)*, Edinburgh, UK.

Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston.

Gill, S. (2013) 'Computer-embedded design: PAIPR prototyping', in Valentine, L. (Ed.): *Prototype: Craft in the Future Tense*, Bloomsbury.

Gill, S., Loudon, G., Woolley, A., Hare, J., Walker, D. and Dix, A. et al. (2008) 'Rapid development of tangible interactive appliances: achieving the fidelity/time balance', *International Journal of Arts and Technology*, Vol. 3, Nos. 3–4, pp.309–331.

Gould, J.D. and Lewis, C. (1983) 'Designing for usability – key principles and what designers think', *CHI '83*, pp.50–53, ACM Press, Boston.

Hare, J., Gill, S., Loudon, G. and Lewis, A. (2013) 'The effect of physicality on low fidelity interactive prototyping for design practice', *Human-Computer Interaction – INTERACT 2013*, pp.495–510, Springer, Berlin Heidelberg, Cape Town.

Hare, J., Gill, S., Loudon, G., Ramduny-Ellis, D. and Dix, A. (2009) 'Physical fidelity: exploring the importance of physicality on physical-digital conceptual prototyping', *Human-Computer Interaction – INTERACT 2009*, pp.217–230, Springer, Berlin Heidelberg, Uppsala.

Haugeland, J. (1998) 'Mind embodied and embedded', in Haugeland, J. (Ed.): *Having Thought*, Harvard University Press, Cambridge, MA.

Lim, Y., Stolterman, E. and Tenenberg, J. (2008) 'The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas', *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 15, No. 2, pp.7:0–7:27.

Maulsby, D., Greenberg, S. and Mander, R. (1993) 'Prototyping an intelligent agent through Wizard of Oz', *Interchi 93*, pp.277–284.

McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B. and Vera, A. (2006) 'Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success', *CHI '06 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp.1233–1242, ACM.

McGrenere, J. and Ho, W. (2000) 'Affordances: clarifying and evolving a concept', *Graphics Interface 2000*, Montreal.

Merleau-Ponty, M. (1945) *Phenomenology in Perception*, Trans., D.A. Landes, Paris.

Molich, R. and Dumas, J.S. (2008) 'Comparative usability evaluation (CUE-4)', *Behaviour & Information Technology*, Vol. 27, No. 3, pp.263–281.

Molich, R. and Jeffries, R. (2003) 'Comparative expert reviews', *CHI 2003*, pp.1060–1061, ACM, Ft. Lauderdale, Florida, USA.

Nielsen, J. (1990) 'Paper versus computer implementation as mockup scenarios for heuristic evaluation', *Interact '90*, pp.315–320, Cambridge, UK.

Nielsen, J. (1993) *Usability Engineering*, Morgan Kaufmann Publishers, San Francisco, CA.

Nilsson, J. and Siponen, J. (2006) 'Challenging the HCI concept of fidelity by positioning Ozlab prototypes', *Advances in Information Systems Development*, pp.349–360, Springer, USA.

Norman, D. (1998) *The Design of Everyday Things*, Basic Books, New York.

Overbeeke, K.A. (2013) 'Industrial design', in Soegaard, M.A. (Ed.): *The Encyclopedia of Human-Computer Interaction*, 2nd ed., The Interaction Design Foundation, Aarhus, Denmark.

Reeves, S. (2006) 'Physicality, spatial configuration and computational objects', *First International Workshop on Physicality*.

Rubin, J. and Chisnell, D. (2008) *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*, John Wiley & Sons, Indianapolis, Indiana.

Rudd, J., Stern, K. and Isensee, S. (1996) 'Low vs. high-fidelity prototyping debate', *Interactions*, Vol. 3, No. 1, pp.76–85.

Sauer, J., Seibel, K. and Ruttinger, B. (2010) 'The influence of user expertise and prototype fidelity in usability tests', *Applied Ergonomics*, "Vol. 41, No. 1, pp.130–140.

Sefelin, R., Tscheligi, M. and Giller, V. (2003) 'Paper prototyping – what is it good for?: A comparison of paper- and computer-based low-fidelity prototyping', *CHI '03: New Horizons, Short Talk*, pp.778–779), ACM.

Sharp, H., Rogers, Y. and Preece, J. (2007) *Interaction Design: Beyond Human-Computer Interaction*, 2nd ed., John Wiley & Sons Ltd., Chichester, UK.

Snyder, C. (2003) *Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces*, Morgan Kaufmann, San Francisco, CA.

Virzi, R., Sokolov, J. and Karis, D. (1996) 'Usability problem identification using both low and high-fidelity prototypes', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground*, ACM, Vancouver, British Columbia, Canada.

Virzi, R.A. (1989) 'What can you learn from a low-fidelity prototype?', *Proceedings of the Human Factors Society 33rd Annual Meeting*, pp.224–228, Sage Publications, Santa Monica, CA.