Interactive Television for Young Children: Developing Design Principles

Ana Vitoria Joly Hulshof

A thesis submitted in partial fulfilment of the requirements of the University of Brighton for the degree of Doctor of Philosophy

September 2010

University of Brighton

Abstract

The research reported in this thesis investigates preschoolers' interactions with interactive television applications. The study involved the development of an electronic programme guide prototype and the empirical evaluation thereof. There were three main aims. The first aim was to analyse children's interactions and illustrate them in a framework to further understanding of the way preschoolers interact with the television. The second aim was to contribute design principles for preschool interactive television and the third aim was to refine methods and add to the knowledge of design and evaluation techniques involving young children.

This research, which involved design and evaluation phases, was carried out with children in Brazil and in the United Kingdom aged between three and four. Children participated actively as informants and were asked for input at various stages of the project. Their participation during design activities and evaluation sessions was crucial to the constitution of the framework, the development of design principles and the refinement of methods for working with preschoolers.

The results revealed that young children's interactions with interactive television are influenced by: the children's age, motor skills, country context, media and device use and knowledge. These factors along with the input device used for interaction and characteristics of the interactive television application determine the number of accomplishable tasks, the time taken to accomplish the tasks, the complexity of the tasks that can be accomplished, the hints and assistance as well as interactions needed to accomplish the tasks. The combination of such aspects is then reflected in the user experience. The main issues that occurred during the interaction with the prototype application are documented, and a list of design principles is presented to assist in the design and evaluation of interfaces that meet the needs, capabilities and interests of young children. Additionally, existing methods for design and evaluation were refined throughout the project and novel activities developed. The intention is that these contributions will be useful to designers and researchers of interactive television and those in the field of interaction design for children.

Brief Table of Contents

Detailed Table of Contents

List of Tables

List of Figures

Personal Preface

In the early 1980s when I was three years old the amount of media content available specifically for preschoolers was scarce. At home, my parents believed television would interfere with my development so I was not allowed to watch television until the age of four. Even after that I could only watch a couple of educational TV programmes and all my media consumption was strongly mediated. Today, while I believe excessive television viewing and inappropriate content may have negative effects on young children, programmes produced specially for preschoolers can contribute to their development and an element of active participation could certainly enhance the benefits.

I have been fascinated by the potential of interactive television since I started to research the topic during my first degree in 1999. After graduating I worked on a children's programme aimed at preschoolers, allowing them to interact by sending pictures, videos and drawings through the post. I was really intrigued with the tremendous amount of content received from such young viewers and interested in the possibilities that digital television technology could offer to this audience.

In my opinion, children should not be motivated to watch more television than they already do, but appropriate content may benefit them and interactivity can enrich their experience during the time they already spend with the media; prohibition then is not necessary or justifiable.

Acknowledgements

First and foremost I would like to thank my supervisors Lyn Pemberton and Richard Griffiths. Thank you for your guidance and support during all the stages of the research and for always being so generous and accessible.

This work could not be completed without the cooperation of the nurseries in the United Kingdom and Brazil. Thanks to children, parents and staff from: One World, Bright Start, Kipling Lions, Roundabout, Tarnerland and Wishing Tree Nurseries (Brighton); Trenzinho Feliz and EMEI Clycie Mendes Carneiro (São Paulo), the Childcare at Lara and Barnaby Bright (London).

I am grateful to Renata Yumi Shimabukuru for her assistance with the card sorting activities and all the inspiring discussions about children and technology.

I also would like to thank Richard Castle and Myra Wiseman for their help with the statistics.

Thanks to Aimee Lam for the narration of the prototype and to Micah Rosenkind for the source code of the prototype 3D wheel.

Many thanks to the experts who contributed with the prototype evaluation: Judith Masthoff, Jan Koeppen, Marcus Winter, Heloisa Candello, Jean Gabriel Ganascia, Angela Bah, Almir Almas, Susanne Sperrring, Julia Maly, Diana Xu, Mark Rice, Mark Springett, Douglas Carvalho, Nisha Weeramanthri, Junia Anacleto and her study group on human-computer interaction at the Federal University of São Carlos (Brazil) and also Andrew Turk for his valuable comments.

Finally, I would like to thank my examiners, George Lekakos and Avril Loveless, for the valuable feedback, their comments have significantly improved this thesis.

Funding

The work in this thesis was made possible by funding from the Overseas Research Students Awards Scheme (ORSAS) and University of Brighton scholarships, as well as the Faculty of Management and Information Sciences Research Fund, their support is gratefully acknowledged.

Dedication

This thesis is dedicated to my husband, Eduardo Hulshof, who illustrated and programmed the prototype and was part of every single step of this journey.

Declaration

I declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the original work of the author. The thesis has not been previously submitted to this or any other university for a degree, and does not incorporate any material already submitted for a degree.

Signed

Dated

Chapter 1. Introduction

1.1 Introduction

This thesis is a contribution to the research field of interactive television (iTV), human computer interaction (HCI) and its young subfield of interaction design for children (IDC), which is concerned with the design, evaluation and implementation of interactive systems for young users. This work is focused on the design of interactive television interfaces for preschool children.

This chapter provides a brief overview of the context of the thesis (Sections 1.2 to 1.4), the aims and objectives of the research (Section 1.5), followed by an outline of the thesis structure (Section 1.6).

The effects television has on children are controversial, but not the focus of this thesis. Nevertheless, the following section acknowledges where this work stands in this debate. The research is not focused on television history either, but a brief review of children's television and the nature of children's programmes are included in Chapter 2 in order to place the work of the thesis in context.

1.2 Children and Television

From the beginning of television there were already concerns relating to the impact of the medium on children (Maccoby, 1951). During the past years, research has reported that screen entertainment could affect young children's development and guidelines were released initially recommending that children under the age of two should not watch television (Zimmerman and Christakis, 2005). More recently, a British study suggested that children under three years of age should have no screen exposure (Sigman, 2007). There is extensive research about the types of damage screen media could cause to children's development, such as obesity (Robinson, 2001), attention disorder (Christakis et al., 2004) and aggression (Eron et al., 1972).

On the other hand, some researchers have other perspectives. There are several studies that highlight the benefits of television for cognitive and social development (Baydar et al., 2008, Close, 2004, Davies, 1989, Gentzkow and Shapiro, 2008).

The fact is, today's children have increasing access to a range of electronic media and different content is specially produced for each particular age group. Denying preschoolers access to this content would prevent them experiencing the legitimate benefit of media. Among a variety of media devices and content such as computer games, the internet and portable electronic devices, television still appears to be a secure part of children's cultural ‘diet’ and it is more important than other media because of its universal accessibility to all classes, ages and types of children (Davies and Thornham, 2007).

1.3 Interaction Design for Children

As previously mentioned, interaction design for children is a subfield of the human computer interaction field, built not only on the foundations of computing research but also on psychology, education studies and theories. Research in this field is concerned with the way children interact with technology. Markopoulos et al. (2008) state that “the term interactive technology suggests that something extra is added to an interaction by the technology. Thus, the technology must operate in such a way that particular inputs result in specific (different) outputs” (Markopoulos et al., 2008, p.22).

There are different types of interactive technology available for preschoolers. Research ranges from studies on the use of computers in preschool settings (Plowman and Stephen, 2005) to design, implementation and evaluation of tangible programming blocks (Wyeth and Wyeth, 2001) and online tools (Burton and Gould, 2005); children’s interactions with robotic pets (Kahn et al., 2006) and interactive tabletop technologies (Mansor et al., 2008); as well as opportunities and challenges for preschoolers with regard to motion-based video game consoles such as the Nintendo Wii (Bryant et al., 2008). As may be noticed from the examples above, there are several types of interactive technologies for young children, some focused on education, others on entertainment. Read (2005) includes an additional genre often omitted and divides interactive products for children into three categories, education, entertainment, and enabling products, to include dedicated word processors, web browsers and communication tools.

In this young field there is a growing body of research focused on both the design of interactive technology for children and evaluation of interactive technology. In the design process, children may play different roles, from design partners to users. Their participation may vary from being part of the entire design cycle, negotiating design decisions with adult designers, to testing existing designs so that their experience can be analysed to benefit future design of interactive technology (Druin, 2002). Evaluation, in its turn, may be focused on usability (Hanna et al., 1999), accessibility (Gibson et al., 2003), learning (Plowman et al., 1999) or enjoyment (Read and MacFarlane, 2006).

In the research presented in this thesis, children participated as informants (discussed in Chapter 3) in the design and evaluation of an enabling interactive television application. During this process, the usability, accessibility, user experience and preference for input devices were measured, along with children's comprehension of the interface.

Contribution to knowledge: interaction design for children.

With respect to interaction design for children, the main contribution of the thesis is the exploration of a variety of techniques and development of principles for design (in Chapter 4) and evaluation sessions (in Chapter 5) involving young children.

1.4 Interactive Television

Children's television has always been interactive. Around the globe, viewers since the early ages of TV have been invited to dance, sing along and make arts and crafts. In the 1950s, a US programme pushed the boundaries and asked children to interact by drawing *on the television screen* using a special kit to help the main character with what he needed during his adventures. Broadcast in the United States on Saturday mornings, *Winky Dink and You* is considered by some researchers to be the first truly interactive television (iTV) programme (Gawlinski, 2003).

There are several definitions for interactive television. The definition may be strictly related to programmes that are digital and make use of a return path to establish a dialogue between the viewers and the broadcaster. Alternatively, it could include interactivity through use of the telephone, letters or simply by asking the audience to perform activities as described above. In this thesis, interactive television is defined as programmes, applications and services that the user interacts with using a device connected or directly linked to the screen (for example, remote control, mouse). In this case the user is able to interact and alter the audiovisual content being displayed. The interactivity may take place locally, on the set top box or another type of receiver or via return path, and it is not limited to traditional TV.

There are different types of interactive television applications such as teletext-style services, in which viewers can find additional information relevant to a channel. There are walled gardens including a variety of regulated content and services, like games, news, email, shopping - all available under one umbrella. Internet on television, as the name suggests, provides access to internet content through the TV. In enhanced television, overlays of text and graphics are added to programmes providing viewers the opportunity to interact with the television programme while watching it. Additionally, there are also video-on-demand and personal video recorders, applications that provide viewers with access to programmes, films and events at any time, allowing them to pause, rewind and forward the content. Finally, electronic programme guides (EPGs) are “one of the most useful and important types of interactive television" (Gawlinski, 2003, p.7). They display the selection of content available, allowing viewers to search, browse and choose programmes to watch.

In this research, the electronic programme guide was the interactive television application chosen for the prototype application to be designed and evaluated. The process of selection of the application is detailed in Chapter 3 (Section 3.4.2).

Currently there are very few major studies on interactive television for children. Chorianopoulos and Lekakos (2007) explore the characteristics of interactive TV that facilitate education and play focusing on a wider and older age group. Hynd (2006) examines comprehension, attention and enjoyment of young children viewing different types of interactive TV programmes compared with those of children viewing non-interactive versions of the programmes. Finally, Weeramanthri (2008) analyses preschool children's responses to interactive television, particularly in relation to key pressing behaviour.

Contribution to knowledge: interactive television.

The research presented in this thesis offers an insight into electronic programme guides for young users, and, more importantly, adds to the knowledge on interactive television for children by providing empirical evidence that preschool children are able to interact with iTV applications on their own. Additional significant contributions include a discussion of the main issues that interfere with children's interactions and identification of design principles for interactive television for children (summarised in Chapter 6).

1.5 Thesis Project

This qualitative research, in which preschoolers are informants, is not focused on the characteristics or effects of the media but on obtaining empirical evidence of how three and four year old children, the youngest users allowed by the medical community, interact with the television, considering them as viewers not learners.

More specifically, the aims of the thesis are to:

- Analyse the complexities and details of these interactions and illustrate them in a framework to aid with further understanding of the way preschoolers interact with the television.

- Contribute design principles for preschool interactive television.

- Refine methods and add to the knowledge of design and evaluation techniques involving young children.

In order to achieve these aims, participants' contributions during design activities were combined with the data gathered during evaluation sessions to formulate a framework of preschoolers’ interactions with iTV applications. The implications of the design for children’s interactions were examined and design principles for interactive television for young users generated to assist in the design of interfaces to meet the needs, capabilities and preferences of preschool children. During the process, techniques to work with young children in technology design and evaluation were developed and refined.

Behind the work of this thesis is the assumption that preschool children are able to interact with iTV applications using the remote control and mouse. In particular, the thesis identifies the main factors that interfere with children’s interactions (illustrated in the framework in Chapter 6); provides a list of design principles (summarised in Chapter 6); and discusses methods to work with young children in technology design (in Chapter 4) and evaluation (in Chapter 5).

1.6 Conclusion

This chapter has placed the work of the thesis in context and highlighted the main research themes.

The thesis is structured in seven parts. This introduction is followed in Chapter 2 by coverage of the background literature, containing an overview of children and television, interaction design for children and interactive television. Chapter 3 presents the research design and methodology. Chapter 4 includes a description of the design activities and interpretation and implementation of the results into the prototype. Chapter 5 describes the evaluation sessions carried out to investigate children's interactions with the iTV prototype. Chapter 6 discusses and presents design principles and finally, conclusions and suggestions for future work are presented in Chapter 7.

Chapter 2. Relevant Research

2.1 Introduction

This chapter is divided into four sections. This introduction is followed by Section 2.2 with children's television context, a brief history of interactive television and an overview of interactive television for children. Section 2.3 contains the process of gathering requirements for the prototype application without children, comprising of a review of the literature and analysis of existing applications. The conclusions are then presented in Section 2.4.

2.2 Context

2.2.1 Children's Television

This section looks at children's television context in the United Kingdom (UK) and Brazil, countries in which the research was conducted, and also includes some data from the United States that was found relevant to the context presented. The section comprises a concise history of children's television, other issues, such as media effects, are not analysed.

In the 1930's, the world's first regular high definition television service began transmitting in Britain and much effort was expended by the BBC's producers to determine the type of programmes which would prove attractive to viewers (Burns, 1998). Before the Second World War there were programmes of interest to children, but television had no committed children's hour (Oswell, 2002).

Children's television is considered to have officially started in the United Kingdom in 1946 when the BBC launched the programme *For the Children* a ten minute long series broadcast on Sunday afternoons (Buckingham et al., 1999). In 1950 *Watch with Mother* innovated by combining education and entertainment in a programme dedicated specifically to preschoolers (Home, 1993).

After the war the number of programmes dedicated to the child audience started to increase and to define their space in the television schedule. In this context, strong age-related boundaries between programmes were tied with the schedule and supported parents in ensuring children watch content suitable for them, according to its allotted time slot. “Continuity and scheduling were central to the guiding process, and the clearest signpost was built into the fundamental structure of the schedule. Under-fives were given their own separate mid-afternoon slot” (Buckingham et al., 1999). Subsequently, the advent of the UK channel BBC2 in 1964 gave room to new children's productions like the daily preschool series *Play School*, and in the late 1960s the output of BBC Children's department was between nine and ten and a half hours per week (Home, 1993).

Internationally, children's TV also started around the same period. In the United States, television programmes for children began to be aired nationally in the late 1940s when the NBC success *Howdy Doody Show* became one of the first children's programmes to be broadcast five days a week. Shortly thereafter, the Americans recognised the potential and advantages of the medium for the younger audience as an educational tool, motivating the consumer market. “In the 1950s, children's programs and the benefits that television could presumably bring to the family were highly touted selling points for television sets” (Alexander, 2004, p.502).

In Brazil the first children’s programme was broadcast in 1955; *Clube do Guri*, originated from a radio programme, featured prodigy children who sang, recited poems and played several instruments (Pereira, 2002).

Until the 1980s, there were no significant changes in the children's television scenario. The amount of content produced increased and most countries broadcast children's educational and/or entertaining programmes after school, on Saturday mornings and Sunday teatimes. In the late 1980s, cable television and the VCR provided an extreme flexibility to the schedule. Video recorders offered the chance to have children's favourite’s programmes on tape to be watched at any time and channels dedicated to children broadcast content for them with no interruption.

The multi-channel era in the United Kingdom was transformed with the arrival of digital television, providing access to better quality of image and sound together with more channels. In 2002, children who before could only have access to limited content through terrestrial TV or most American content through cable or satellite, were then benefited with the launch of two BBC channels, CBeebies and CBBC.

In Brazil terrestrial digital television was implemented in 2007 but does not yet offer a channel especially for children. Brazilians have no free access to children’s channels thus far, but they may watch dedicated channels broadcasting imported content and a Brazilian channel, TV Rá Tim Bum, via cable and satellite.

Over the past few years, the number of channels for children has multiplied and each network has currently several channels aimed specifically at each age group. The traditional way in which television schedule related to children's routines no longer exists as the children’s TV context moves from children's hours to children's channels. Lury (2002) states that channels become “places” and describes this transition as “from *when* to *where*”, according to her this shift means children have to orientate themselves spatially rather than temporally, she explains:

> “That channels and schedules have become brands is therefore symptomatic of a changing media landscape, but it also suggests that the temporal positioning of programmes –

> when something is 'on' – is increasingly being displaced by spatial concerns – where something is 'on'" (Lury, 2002).

At the present, this context has changed with a significant increase in the on-demand content available for children through the set top box and the internet. Children's niche channels still exist but they are now complemented by hundreds of films, cartoons and complete series available at any time.

Children's on-demand content on the internet is available on several video hosting websites, such as YouTube, but also through video players that are part of each TV channel's website. In this sense, children still orientate themselves "spatially", the website being equivalent to a channel, the "place" to find a specific programme.

The organization of the content available on-demand via the set top box is, however, not yet well defined. Children's programmes are merged into extensive textual menus in which children have to choose a category, then a subcategory, then the programme and finally the episode they wish to watch. Different ways to find on-demand programmes include choosing the programme from a list organized alphabetically and by the date it has been broadcast. These are certainly not relevant for preschoolers. There is the option to choose the name of the channel as a category, which could be related to the "place", but the process of finding the content also consists of a complex navigation through a textual interface.

Figure 1. Electronic programme guide for terrestrial digital TV Samsung©

As a result, it can be said that children can still find audiovisual content in a "place" that can be a website or a category in the on-demand menu. Nevertheless, the main difference between the multi-channel and the on-demand context is that children can choose the specific programme they want to watch but to find it they have to undertake, on both website and digital television, a much more complex interaction process than the previous changing channels'.

Accordingly, the landscape of children's TV moves from *when* to *where* to *what*. Besides being able to choose the type of content they wish to watch, by choosing a particular channel, children can also choose the specific programme they want to view (Figure 2). Multi-channel TV has been complemented by the on-demand facilities in cross-platform environments offering children choices to watch anything, anytime, anywhere.

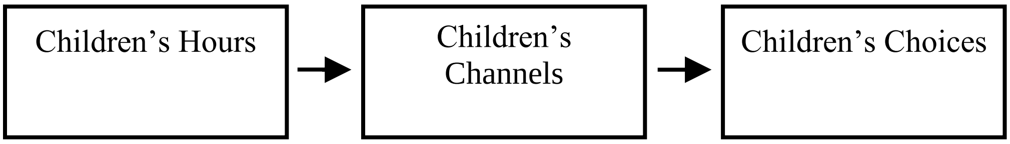


Figure 2. Evolution of children's television

Interactive applications involving the television screen such as video-on-demand, which before had no impact on children, "except as part of futuristic fantasies" (Livingstone and Bovill, 2000, p.20) are today a reality. In the United States, Sesame Workshop, PBS, HIT Entertainment and Comcast created a 24-hour channel for preschoolers with on-demand service to provide children and their parents a selection of programmes no matter what time of the day (Knell, 2006). And in the United Kingdom, the Office of Communications (Ofcom) reports (2008) that on-demand content is now widely available, major broadcasters such as the BBC and Channel 4 are offering the service over the internet and cable and younger consumers are showing a growing interest in accessing it.

One can argue that those changes in context do not actually add more choice since most of the content producers are the same as they were in the first stages of children's TV. But it surely offers children the possibility of watching a great range of programmes produced over the years independent of the time of the day or the channel they were originally broadcast.

2.2.1.1 The Nature of Children's Programmes

The nature of children's television programmes drew on an established set of genres developed within children's radio in the 1920s and 1930s. As with children's radio, TV programmes were constructed with a general ethos of participation for the child and "to connect the child to an external world in an active form of citizenship" (Oswell, 2002). In these early days, the collection of genres that composed children's television included story-telling, current affairs and children's music (Home, 1993). It attempted to provide a "microcosm of adult schedules" (Davies, 2001).

Davies and Thornham (2007) underline the importance of broadcasting a range of genres, given that children's preferences are very divergent and the benefits of pleasing diverse audiences come under three broad headings: learning, socialisation and citizenship, and thirdly personal fulfilment and identity. Specially focused on learning and socialization, in the United States, following several early educational programmes, such as *Ding Dong* and *Romper Room*, in the 1960s the American Public Broadcasting Service (PBS) aired *Sesame Street* as "an experiment in using television for social change" (Pecora et al., 2007). The pro-social *Sesame Street* is the most studied series in the history of television (Fisch and Truglio, 2001) and demonstrated over the years that television can teach preschoolers educational concepts. Originally, its educational goals were divided into three categories: cognitive, affective and physical goals (Stein, 1979). Research has confirmed that the programme assists children in improving their vocabulary (Rice et al., 1990), acquiring reading skills (Wright et al., 2003) and counting skills (Fisch and Truglio, 2001).

Sesame Street excellence was internationally recognised. There were, however, some controversies. The BBC justified not buying *Sesame Street* because an hour programme was too long for children under five to keep still watching TV and they preferred to encourage children to think and understand the reasons behind statements instead of learning by repetition (Home, 1993). Gwen Dunn (1977), along with other researchers such as Grant Noble (1975), also criticised *Sesame Street* and were opposed to television's teaching stance. They suggested that TV programmes should provide children what schools cannot offer, leaving the teaching (skills) to teachers. For Dunn “television can seal things in a child's mind if they are repeated often enough but this is a long way from the acquisition and use of a skill” (Dunn, 1977). *Sesame Street* was broadcast in several countries including Brazil, where fifty per cent of each episode was American content dubbed and the other fifty per cent was content produced with a Brazilian cast (Carneiro, 1999). The programme had a big audience in Brazil but there were criticisms about its American origin and its pedagogic aspects (Caparelli, 1982).

Today, there is still “a remarkable divergence between the dominant educational philosophies on either side of the Atlantic” (Buckingham, 2002). Buckingham compares the child-centred 'progressivist' approach of the British programme *Teletubbies* with the most didactic approach of the American programme *Barney*.

Along with the educational theory debate there is the controversy between educational and entertaining content. Education, especially for young children, has to be entertaining to engage them; and both education and entertainment play an important role in their development.

It is arguable, though not the focus of this thesis, that children may benefit more from entertainment than educational content or from a didactic rather than a less instructional approach. But not surprisingly, research had shown that in any scenario children can benefit more from television programmes if there is social interaction, adult participation and interaction asking questions and providing feedback (Reiser et al., 1984), clarifying content and adjusting it according to the child's needs. This is in line with Vygotsky's theory (further discussed in Section 2.3.1.2.2) that "cognitive development is essentially a social process" (Schaffer, 2004). Vygotsky, one of the most influential psychologists, identified a period that fell between two levels of development in which children could perform at a higher developmental level with assistance of a more skilled person (Vygotsky, 1978).

It is important for parents and guardians to help children both to learn from television programmes as well as to acquire a critical view of the content watched so they can benefit more from what the TV offers. However, regardless of whether the child did not understand a word or a segment of a traditional linear programme it will continue from beginning to end. On the other hand, in the present scenario, in which children have to interact to find the programme they want to view and can also interact with services and games, if they cannot achieve the development level required to perform the task they will be stuck.

Many parents are grateful that their children are learning from TV, but it appears that the primary reason they choose to bring media into their children's life is not because of the educational factor, but because of the practical benefits it offers: uninterrupted time for chores, some quiet time, or even just an opportunity to watch their own favourite programmes (Rideout and Hamel, 2006, p.5). This means most children watch television unaccompanied and would not have the Vygotskian scaffold a co-viewing experience could offer. For this reason, "there seems to be an advantage in making technology play a more social role in supporting children's learning" (Ryokai et al., 2003). Media can play the role of a more experienced user that through instruction and modelling provides aid in intellectual growth to a less experienced user.

In this scenario, the viewer becomes a user. In other words, children watch and interact and their interactions change the content displayed on the screen. Children's television production meets interaction design for children in order to provide the structure for children's interactions and enhance media's benefits. This accords with a reality in which asking children to *Watch with Mother* is not always realistically practicable.

2.2.2 A Brief History of Interactive Television

As mentioned in the introduction, some researchers (Carey, 1996, Lu and Frye, 1992, Lu, 2005, Gawlinski, 2003) consider the first interactive television programme *Winky Dink and You*. Broadcast in the 1950's, it motivated children to help a cartoon character out of a jam by drawing objects such as a bridge, a rope or a ladder, on a transparent sheet fixed on the television screen. In this thesis *Winky Dink and You* is not considered an interactive television programme because children's interactions would not actually influence the content. The character would behave the exact same way and the story would unfold accordingly despite the fact that children draw a bridge for him to cross a river, a submarine, or did not draw anything at all. Children's different inputs could affect their experience while watching the programme, if, for instance, they danced or sang along, but did not result in specific outputs.

In the early seventies, Ceefax was launched by the BBC in the United Kingdom. The teletext system, still broadcast today, takes advantage of unused parts of the broadcast signal to transmit text and simple graphics that the viewer can access by typing page numbers on the remote control (Gawlinski, 2003). This would then be in fact the first interactive television application, according to the definition used in this thesis, since viewers' interactions though the remote control results in specific content being displayed on the screen. It is worth highlighting that changing channels, as well as other aspects such as volume and contrast, is not considered interactivity in this perspective.

Following Ceefax there were a series of interactive television experiences such as the QUBE, in the United States, in which viewers equipped with a set top box could participate in game shows, choose sport events, order pay per view content and participate in opinion polls. This service, broadcast in the United States, failed due to its high costs for the viewers and for the cable operator (Jensen, 2008). In the eighties, two other American services were shut down because they were not commercially viable; Index and Time Teletext were experiments with teletext and videotext that also failed (Gawlinski, 2003).

In the 1990s' the internet met the television. There were the first examples of two-screen interactive television in which channels like MTV in the US and Channel 4 in the UK display chat rooms onscreen during television programmes (Gawlinski, 2003). Later in this decade, the idea was to strengthen this relationship so that the universal market spread of the television would be combined with the web content (Jensen, 2008), to provide a one-screen interactive internet television.

The transition from analogue to digital, in the late 90s', made internet on the television feasible. Compared with Ceefax, a much larger amount of data could be sent along with video and audio. In addition, viewers could send information back to the broadcasters via a return path establishing a direct interaction in which the interactive service supplier is able to act on the request sent by the viewer and send information out (Gawlinski, 2003). At this stage in the United Kingdom interactivity via the red button arrived, giving viewers access to a range of services and information by pressing the red button on the remote control and enabling the viewer to interact with the broadcasters directly without picking up a phone or using a computer (Katz, 2004).

The paths of the internet and interactive television have crossed several times over the past twenty years. Initially there were some issues with the idea of giving access to millions of pages and services available on the internet through television. This was because television and computer displays had different resolutions that required specific colours, fonts and screen layouts (Herigstad and Wichansky, 1998). Another factor that differed in both media was that viewers watched television at a further distance from the screen than they used for the computer. Television viewing was therefore characterized as a "lean back" experience opposed to the active and engaged "lean forward" to interact with the computer (Masthoff and Pemberton, 2005). This scenario has changed since television screens have gained resolution with the spread of high definition TV sets (HDTV). Television is increasingly viewed through computers that act as facilitators for content acquisition and replay (Barkhuus, 2009) and web services have been developed as widgets for TV (Shin et al., 2008). The screen resolution and viewer/user behaviour when interacting with the television and computer might be gradually more similar, but there is still a difference in the input devices used for interaction. Today, the remote control is still the most common device used to interact with the TV and set top box, while television through the computer is usually controlled with the mouse and keyboard, although the remote may be used. These are currently the most popular input devices, but there is a range of additional devices used to interact with television content on TV and/or computer. Game consoles offer access to video and the interaction occurs via joysticks and other controllers. Touch screens and multi touch surfaces are not as widespread. Nevertheless they can be used to interact with television as well, and, in the near future, brain computer interfaces (Nijholt and Tan, 2008) may be used to control TV content. The navigation with each input device, however, is different and for this reason the integration of the "lean forward" and "lean back" experiences presents a challenge for designing flexible interfaces in which interaction is straightforward with any input device, providing viewers a good television experience on their sofa, on their desk or elsewhere.

2.2.3 An Overview of Interactive Television for Children

As previously mentioned, research on interactive television for children is, to date, scarce. Thus, in this section, are discussed the few studies focused on iTV for children in which the interaction takes place via the conventional remote control. In addition, an overview of different experiments on children's interactions with video content via alternative input devices, not specified as television, is presented.

Chorianopoulos and Lekakos (2007) investigate opportunities offered by interactive television for education and entertainment. The authors state that new issues need to be considered to successfully support learning and playability. Hence, the traditional HCI considerations should be extended and emphasis given to engagement and fun instead of focusing only on usability and usefulness (Chorianopoulos and Lekakos, 2007).

Hynd (2006) investigates which different types of interactivity provide benefits to young children. The author highlights that: only specific models of interactivity result in higher comprehension, attention and enjoyment. Programmes allowing children to repeat sections and those with increased viewer participation result in higher attention and comprehension while the possibility for customisation did not result in any benefits and narrative choices were associated with lower attention, comprehension and enjoyment (Hynd, 2006).

Weeramanthri (2008) analyses children's responses to interactive television prototypes and explores the relationship between key pressing behaviour and levels of motivation, mastery and breakdown. The author concludes that interactive television affords participation by preschool children and indentifies as factors affecting preschool children's responses to interactive television:

- The nature of the interactive task
- The level o difficulty of the interactive task
- Prior experience with interactive media
- Gender

It may be noticed, from the very limited published research on interactive television for children, that there are contradictions about which type of interaction should be provided. Weeramanthri (2008) suggests that customization and choice is an appropriate modality for preschool-aged children viewing interactive television. Hynd (2006), on the other hand, states that customization does not result in any benefits and choices are negatively associated with attention, comprehension and enjoyment. There is no further literature concerning content, navigation and screen design specifically for children's iTV.

Thus, it is worth providing an overview of a range of experiments and studies on children's interactions with different types of video content through different input devices that are not specified as television but may be similar to TV content. Palenque, for instance, was an interactive multimedia digital video interactive prototype based on a television programme in which 8 to 14 years old children could navigate around the video environment and explore a virtual ancient Maya site (Wilson, 1988). Plowman (1991) investigated the use of interactive video in secondary schools and identified design principles by analysing children's interactions from a learner's perspective. In order to interact with Palenque children had to use a joystick, while in Plowman's study the trackball and keyboard were used as input devices. Both studies were, however, aimed at an older age group, for preschoolers a common input device used to interact with video content are toys (Luckin et al., 2003b, Strommen, 1998, Tuchscherer, 1988).

It has been debated since the early years of children's television the strong connection between television programmes and an enormous range of toys, in which programmes that feature characters available as toys may act as full length commercials target at children (Gunter and Furnham, 1998). Study in this field is characterized as polemic and interesting, but it diverges from the central theme of this research. Thus, as well as advertising, it will not be discussed in this thesis. Addressed instead are toys and devices that allow the child to control or obtain responses from on-screen content bridging the gap between linear and interactive television.

In the 1980s' TV interactive toys gave children the opportunity to join characters in futuristic battles. *Captain Power* was a programme broadcast in the United States and Canada in which children could use Mattel's power jets and spaceships to shoot at the television and the television would then activate the toy. Children had to shoot at enemies on-screen and defend themselves by firing at on-screen projectile. They would score points when targets were hit and if they missed their toy would take an enemy hit and suffer imaginative damage effects resulting in losing points. *Saber Rider and the Star Sheriffs* was another programme broadcast in the 80s', in the United States among other countries, in which children could purchase interactive vehicles and action figures to shoot at the enemy on the screen. The score was kept on the toy, accompanied by lights and sounds from the weapon, and the communication between the TV programme and the toy was via infrared waves from a decoder fit onto the TV to the toy. These programmes were heavily criticised for encouraging children to role play aggressors and defenders in battle scenes and involve them in participatory violence encouraging aggressive behaviour (Tuchscherer, 1988).

Toys children could use to interact with the television and the computer for educational purposes have been also criticised. ActiMates Barney, for instance, developed by Microsoft was an animated plush doll for two to five year olds that could be used as a freestanding toy and via a radio link interacted with the software on the computer and videotapes played on the TV (Strommen, 1998). The toy was aimed at engaging children in learning interaction, but it raised some concerns because it told the child what to do instead of supporting creativity and giving the child the control, and some researchers believed it could generate high degree of emotional connection offering an alternative to real emotional connection with another person (Fogg et al., 1999).

Additional ActiMates digital soft toys such as Arthur and D.W. were also released by Microsoft and combined with compatible software running on the computer to comment on children's interactions, provided feedback and support. The toys' assistance was found to be inadequate and even inappropriate but Luckin et al. (2003b) concluded that such sophisticated systems could use the potential offered by tangible technologies to provide richer learning interactions. Furthermore, it was believed that the toys could stimulate interaction and collaboration among peers, readdress the gender imbalance in the educational use of computers and bridge the gap between formal and informal learning contexts (Luckin et al., 2002).

Microsoft discontinued the Actimates toys in 2000, but today there is still a range of tangible technology available that provide young children control and/or feedback from on-screen content.

Among different types of preschool entertainment MagIQ, aimed at 6 to 36 months children, includes a DVD and a teddy bear. The toy plush animal reacts in real time to on-screen activities in the DVD and encourages interaction between the baby, the toy and the TV content (Hayes, 2008). The toy manufacturers V-Tech and Leap Frog have also produced a range of educational interactive toys that enable children from nine months to nine years of age to control on-screen content with age-appropriate joysticks and devices such as dance mats, microphones and touch sensitive drawing pads.

It is argued that interaction per se does not provide young children a good experience (Hynd, 2006). Successful interactivity should be built upon features of children's television such as opportunities for participation and repetition of the content (Hynd, 2006), should explore scaffolding (Luckin et al., 2003b) and, most importantly, should be developed specifically for preschool children (Plowman and Stephen, 2003).

The examples above indicate that there is significantly more literature addressing children's interactions with video content through alternative input devices than analysing interactive television applications for children. Thus, in order to provide background for the work in this thesis and support the development of the prototype application, extended HCI considerations (Chorianopoulos and Lekakos, 2007) along with IDC research and interactive television studies for adult users will be reviewed in the following section (Section 2.3).

2.3 Requirements

Requirements are statements that determine users' characteristics and the conditions an interactive product will be used in order to specify what it should do and how it should perform. Preece et al. (2002) identified as the most common data gathering techniques used to establish requirements: questionnaires, interviews, workshops or focus groups, observation, and studying documentation. The authors stated that these methods can be combined to support identification of different types of requirements such as functional, data, environmental, user and usability. In this thesis requirements are not divided into types but grouped according to the method they were gathered:

- Requirements from literature (RL)
- Requirements from existing applications (REA)
- Requirements from observing children (RO)
- Requirements from card sorting activities (RCS)
- Requirements from low-tech prototyping sessions (RLTP)
- Requirements from experts' evaluations (RE)
- Requirements from prototype adjustment session (RPA)

In order to establish requirements it is first necessary to identify user needs. This means understand users and the context of use so that the system under development can support them in achieving their goals. From the needs a set of requirements are established to form a foundation to the design. Identifying needs and establishing requirements is an iterative process in which activities inform and refine one another, requirements evolve and develop along with the design.

User participation is essential for the process of the design, as already stressed, but before involving children in this research literature covering child development, existent guidelines for interaction design, especially those for children and interactive television applications were reviewed to determine users' characteristics and their needs in order to define some requirements. This preliminary literature review had also the purpose to assist in the data analysis during the following stages of the research. Subsequent to the review of the literature, existing applications were analysed to provide inspiration and ideas as well as to enlighten functional requirements.

Research suggests that requirement methods may be supported by props such as personas, scenarios and task analysis (Preece et al., 2002). According to Allan Cooper (1999), personas are representations of real people, based on research, that include their goals, behaviours and motivations and allow archetypical users to be present throughout the design cycle. The use of personas evidences specific users and their characteristics that simplify the communication and decision making of a project. Antle (2006) adapted Cooper’s work, proposed child-personas, driven by children’s developmental needs and experiential goals and suggested that personas should be continually validated in conjunction with real children. Scenarios are informal stories that express imagined situations and help in conceptual design (Preece et al., 2002), they can be used along with the personas to establish the context users are while using the system and elicit requirements. In the study reported in this thesis, however, children were involved as informants and the decisions were made based on their feedback instead of on fictional characters and situations. It was decided not to use task analysis either because this technique is mainly concerned with the performance of work (Diaper, 2004) and used to investigate the purpose of interaction, what needs to be done both physically and mentally to achieve a task. Task analysis involve identification of user’s goals broken down into subtasks then grouped together to specify how tasks are performed, such approach would be too formal for simple tasks involved on the EPG prototype and not appropriate for an open-ended task in which children could have unclear goals such as browse content.

Thus, after the review of the literature (Section 2.3.1) and existing applications (Section 2.3.2) the process of gathering requirements continued with several activities involving children. An initial observation was followed by card sorting tasks and low tech prototyping sessions. The prototype was then sent to experts for evaluation focused on re-design and presented to children in prototype adjustments sessions for a final feedback before the evaluation sessions. These activities will be described in Chapter 4 along with the requirements gathered from each method.

2.3.1 Requirements from Literature

2.3.1.1 Introduction

As previously mentioned the process of establishing requirements usually starts by identifying user needs and then producing from the needs a set of requirements to form a basis for the design. In order to understand users and determine some of these needs literature on preschoolers' developmental stage is briefly reviewed so that children's abilities and limitations can be identified.

These needs are then related to existing guidelines for interaction design for children and interactive television in general to indicate additional requirements and assist in the design of the prototype application.

Requirements from children's developmental stage literature and guidelines, heuristics and principles are combined and presented as requirements from literature (RL) in the Appendix A.

2.3.1.2 Children's Developmental Stage

Each child is unique, has an individual growth and development pattern, some singularities of specific users characteristics and skills are underlined in Chapter 5, but in this section a general view of preschoolers' stage of development is briefly presented so children's characteristics and abilities, those that are relevant to the design of technology, can assist in the establishment of requirements.

The requirements derived from this literature are organized into three areas of children's development: physical development, cognitive development, emotional and social development. It is important to note that the holistic perspective is the dominant theme of human development today that analyses development as a whole in which each one of these areas self depend and interact with the others (Shaffer, 1999). Thus the different areas of development are presented separately but, in reality, they are intricately linked.

2.3.1.2.1 Physical Development

Physical development addresses physical growth, including bodily changes and the sequencing of motor skills as well as sensory development in which information is received through the senses.

According to their physical development, three and four year olds have small hands and less developed motor skills. For this reason the first requirement should be to offer them input devices of an appropriate size that are not too big for them to be able to hold with their small hands but are not too small either so that they would not require fine motor skills to interact.

RL1 – Input devices should have an appropriate size/dimension

Regarding their sensory development, preschoolers are able to match and name primary colours (Meggitt, 2007). Thus, the coloured buttons on the remote control can be explored to navigate through the interface given that they would be able to recognise the colour presented on the screen match with the colour button on the remote and press it to interact.

RL2 – Explore the use of coloured buttons

The idea of developmental stages has been heavily criticized for defining a universal standard with fixed boundaries and not including enough stages. Characteristics of children's developmental stage are considered in this section to provide general skills that preschoolers are likely to have, but those may vary significantly and evolve constantly. For this reason the system should be flexible so it can cater to users with different and evolving skills.

RL3 – Provide flexibility

2.3.1.2.2 Cognitive Development

Cognitive development deals with the intellectual growth and the development of thought processes used for recognising, reasoning, knowing and understanding.

Children at this age are in Piaget's preoperational stage of cognitive development in which they are egocentric that means they are able to see things only from their own perspective (Shaffer, 1999). For this reason design and evaluation sessions in which they would have to collaborate could present more challenges than individual activities.

RL4 – Favour design and evaluation sessions in which activities are conducted individually

Three and four year olds begin to understand the classification of objects (Smith et al., 2003). Preschoolers' ability to sort objects into simple categories means they could benefit from a prototype application that take into account their concept of categories.

RL5 – Consider preschoolers' concept of categories

Preschoolers don't read or are not expert readers, have brief attention span and are only capable of holding one thing in memory at a time (Bruckman and Bandlow, 2002). These characteristics indicate that text should be avoided when designing interfaces for this age group. Design and evaluation sessions should be short so they are able to stay focused during the entire session. In addition, the use of sub-menus that requires the user to remember complex hierarchies in order to interact is inappropriate.

RL6 – Avoid text
RL7 – Design and evaluation sessions should be brief
RL8 – Avoid sub-menus

Piaget characterized children as active scientists, and their progression from one stage to the next was the result of children's efforts to *accommodate* any new knowledge with that which they already know, and *assimilate* it into a fuller understanding. This results in an *equilibrium* where knowledge is embedded and part of the child. So development, in Piaget's view, was largely unaffected by interventions of adults in the process. He asserted that the learner has an active and independent role, often selecting activities, while the adult's role is supervisory rather than interactive, so learning is an interactive process between children and their environment (Duffy, 1998, Schaffer, 2004).

Despite the fact of accepting the concept of developmental stages, in Vygotsky's view children's progress from one level to the next did not depend exclusively on development it could instead lead the development. Vygotsky (1978) identified a stage between two levels of development in which children are able to perform at a higher developmental level if they are given guidance, he defined it as the *zone of proximal development*. As mentioned in Chapter 2, this concept has been often used in interaction design for children, so that technology plays the role of a more experienced and through instruction and modelling offers scaffolding to support and extend children's learning (Luckin et al., 2003a, Ryokai et al., 2003). The application should enable children to interact with the prototype, explore and discover, without relying on adults to help. Therefore, it should provide age-appropriate instruction and modelling to assist children achieving the *zone of proximal development* by providing scaffolding and guidance that support users and enable them to move from passive viewers to interactive TV users.

RL9 – Provide scaffold and guidance

2.3.1.2.3 Emotional and Social Development

Emotional development addresses the development of feelings and the way they understand, control and express feelings. Social development concerns the growth of the child's relationships with others.

According to the emotional and social development, three and four year olds generally behave more independently and are strongly self-willed, seek frequent adult approval, show pride in accomplishment, and enjoy jokes and verbal incongruities (Sheridan et al., 1997). Complimenting children during design and evaluation sessions is therefore important. Moreover the interface could include humour.

RL10 – Compliment participants during design and evaluation sessions
RL11 – Add humour

2.3.1.3 Guidelines, Heuristics and Principles

In the previous section a few characteristics of preschoolers' developmental stage were presented for an understanding of the users that elucidated some requirements for the prototype application. In this section existing guidelines, heuristics and principles that have been developed over the years are analysed to contribute to the design and evaluation of technology. Some of the requirements previously presented are reviewed and additional requirements discussed.

The requirements derived from this literature are organized into three areas: requirements for input devices, requirements for the interface design and requirements for working with children in design and evaluation sessions.

2.3.1.3.1 Input Devices

Children's small hands and less developed motor skills suggest that all mouse buttons should have the same functionality (Druin et al., 2001) and a limited number of keys should be used in the remote control.

RL12 – All mouse buttons should have the same functionality
RL13 – A limited number of keys should be used in the remote control

Studies of adult users demonstrated that navigation by colour buttons works well, with colours and labels used carefully to match buttons of the remote only when this correspondence is intended (Daly-Jones, 2003, Eronen and Vuorimaa, 2000). The fact that preschoolers are able to match and name primary colours indicates that the use of colour buttons should also be explored on applications for young users (RL2).

In interactive television applications for children the remote control arrow keys should be used for choosing a menu item and the OK key to confirm the choice. These keys are ideal for menu navigation because they are easily found by touch without the need to look at the remote control (Kunert, 2009).

RL14 – The remote control arrow keys should be used for choosing and the OK key to confirm the choice during navigation

2.3.1.3.2 Interface Design

In the interface design if more than one page is required to present content, in both interactive television and interaction design for children literature, it is recommended that paging should replace scrolling (Gawlinski, 2003, Hutchinson et al., 2006, Kunert and Krömker, 2008). Gawlinski (2003) indicates that it is difficult to operate scrolling pages with the remote control and some viewers may find such computer-based concept confusing, therefore the author suggest presenting pages in series one after the other. Kunert and Krömker (2008) also suggest paging is preferred than scrolling and it should be done using the up and down arrow keys on the remote illustrated with graphical arrows on the screen. Hutchinson (2005) have highlighted that paging is better than scrolling but not ideal for young children because they usually don't find the paging buttons without instructions. Thus, she recommends using large and prominent paging buttons or avoiding paging by decreasing the number of categories in the classification structure.

RL15 – Paging should replace scrolling
RL16 – Large and prominent buttons should be used for paging
RL17 – Decrease the number of categories, when possible, to fit content in only one page

It is recommended that icons should be placed close to each other on the screen, but distanced enough to compensate for inaccuracy in targeting (Chiasson and Gutwin, 2005). This requirement is especially important for interaction using the mouse since using the remote control children do not need to aim at targets on the screen.

RL18 – Place icons close to each other but distanced enough to compensate inaccuracy in targeting

On-screen selectors that children need to move around should be very prominent (Gawlinski, 2003) for them to recognise where it is and make the connection between their interaction with the input device and the movement of the selector on the screen.

RL19 – On-screen selectors should be prominent

The cursor should move logically, it should not jump over links and the most important links should be close to the initial position of the cursor (Ahonen et al., 2008, Kunert, 2009). This requirement is especially important for interaction via remote control or keyboard in which movements are broken into discrete steps and each interaction moves the cursor one step.

RL20 – Cursor should move logically and its initial position should be close to important links

Each icon should have a diameter of at least 64 pixels for children to be able to click (Hourcade et al., 2004) and the target acquisition can be enhanced with *bubble cursors* by resizing cursor's activation area (Grossman and Balakrishnan, 2005).

RL21 – Each icon should have at least 64 pixels diameter
RL22 – Resize cursor activation area to enhance target acquisition

Given young children's limited reading skills, text on interface should be avoided (RL6) to reduce cognitive load (Druin et al., 2001). Icons and real world metaphors should be used instead (Rice and Alm, 2007) associated with simple words (Chiasson and Gutwin, 2005).

RL23 – Use icons and metaphors associated with simple words

Clickable icons should look clickable so that the distinction between clickable and non-clickable items is clear (Gilutz and Nielsen, 2002).

RL24 – Clickable icons should look clickable

Preschoolers appear to be able to hold in memory only one chunk at a time, suggesting that extensive menus and sub-menus should be avoided (RL8), a flattened hierarchy should be used instead (Hutchinson et al., 2006).

RL25 – Use a flattened hierarchy

Children usually focus on the middle of the screen, so important icons should be placed in the middle of the page (Hutchinson et al., 2006).

RL26 – Place important icons in the middle of the page

As in applications for adults, the core functionality needs to be presented consistently and should be available throughout the application (Gawlinski, 2003). Icons should be always visible and placed in the same position on the screen to be recognised rather than recalled (Gilutz and Nielsen, 2002, Nielsen, 2005). Furthermore interaction may be facilitated by using consistency and standards that follow platform conventions (Nielsen, 2005) given that users can apply their previous knowledge to navigate.

RL27 – Make core functionality always visible and present it consistently following platform standards

Randomness and humour can increase the enjoyment of using a system (RL11), but great care must be taken to avoid inappropriate additions (Malone, 1982).

RL28 – Randomness and humour should be added appropriately

As previously mention, children's skills vary and evolve rapidly, so the system should provide flexibility (RL3). In addition, it should be efficient to use so that users are allowed to tailor frequent actions, experienced users can use accelerators, unseen by novices, to speed up the interaction (Nielsen, 2005).

RL29 – The system should be flexible and efficient

According to Piaget's idea of the active and independent child, the interactive environment should then offer room for exploration and self-directed discovery. This research is not focused on learning, and there was no intention to base the prototype on the constructivist concept of *microworld* (Papert, 1980), but Papert's extension of Piaget work was crucial for interaction design for children and some of the characteristics from *microworlds* could be considered to offer inspiration to the prototype application such as: adults should only play the role of supervisors, children should know how to interact with and have control of the prototype with minimum or no training, and the application should match their cognitive stage (Rieber, 1996).

RL30 – Children should be able to use the system without adult assistance
RL31 – Minimum or no training should be needed

An additional important feature for interfaces, according to existing guidelines, is to help users to recognise, diagnose and recover from errors (Nielsen, 2005). Proposed solutions to help with error prevention, recognition and recovery is to make the interface simple (Lu, 2005).

RL32 – Make the interface simple

In addition to the interface simplicity, help should be provided on each page and users should be able to access it by pressing a single button used consistently throughout the application (Kunert, 2009). The instructions on the help section should be presented in an age-appropriate format and should be easy to comprehend and remember (Hanna et al., 1999).

RL33 – Provide help and make it always accessible via the same button
RL34 – Instructions should be age-appropriate, easy to comprehend and remember

Errors may also be prevented by providing instant audible and visible feedback (Steiner and Moher, 1992). Animation and audio on rollovers can help indicating where to find functionality while clarifying what to find on the selected button (Gilutz and Nielsen, 2002). Feedback on selection may prevent users to choose an inappropriate option according to the purpose of their interaction and also adds an entertainment value to the process.

RL35 – Provide audible and visible feedback by adding animation and audio rollovers indicating functionality

An exit option should be always offered (Lu, 2005). It should be assured it is easy for users to exit the service (Ahonen et al., 2008) by providing an emergency exit icon on the screen at all times (Nielsen, 2005).

RL36 – Provide an easy escape route and present its icon on screen at all times

Hynd (2006) identifies that the most promising design features of interactive television applications for preschoolers are participation, such as prompting the viewer to respond questions and point at objects on the screen, and repetition. Thus, these features should be provided when possible.

RL37 – Promote participation
RL38 – Provide opportunity for repetition

Customization and personalization should also be supported when possible, children should be allowed to configure the application according to their personal preferences (Kunert, 2009, Large and Beheshti, 2005).

RL39 – Support customization and personalization

2.3.1.3.3 Working with Children

Children this age are egocentric and not always able to take turns, therefore in design and evaluation sessions activities should be conducted individually (RL4) and when working with a group of children it is important to maintain the child to adult ratio low (Guha et al., 2004).

RL40 – Maintain the child to adult ratio low during sessions with children

It is important during the sessions to take into consideration children's brief attention spam (RL7), compliment and reward participants after task completion (RL10) as well as provide feedback for their actions (Hanna et al., 1997).

RL41 – Provide feedback for participants' actions

Revelle (2003) reported that for preschool age children the remote control is too big and they find it difficult to keep it oriented correctly when using it from a hand-held position, for this reason the author suggested placing the remote on a surface. This requirement regards the use of input device, but it is related to the evaluation sessions in view of the fact that the surface users will use to place their remote controls in the home environment will not usually be decided by technology designers.

RL42 – Provide a surface for participants to place the remote control

2.3.1.4 Conclusions

Following the analysis of the literature children's characteristics were identified that provided evidence of some of their needs assisting on the established of 42 requirements. Initially the aim was to identify requirements for the prototype application, but during the process were also gathered requirements for input devices and for working with young children in design and evaluation sessions. The complete list of requirements derived from the literature is presented in the Appendix A.

2.3.2 Existing Applications

2.3.2.1 Introduction

Having decided that an electronic programme guide would be the application to be developed, following the analysis of the literature (Section 2.3.1), the process of design continued with a brief review of current systems. In the industry such review is usually called competitor analysis and it is aimed at identifying strengths and weaknesses of competing products to assess issues that need to be addressed in order to compete effectively (Nielsen and Mack, 1994). In the work described in this thesis, however, the existing applications were reviewed to provide to the prototype to be developed some functional requirements, indicating what it should do, as well as inspiration and ideas for the system and interface concept.

2.3.2.2 Electronic Programme Guides

An electronic programme guide is an application that allows users to search and browse video content. This means users may look for a specific video within the content available or they may explore the content in an open-ended way.

Today children may search for a particular programme using EPGs via digital television or via the internet. There are a number of ways that the search may be conducted. A list of programmes available for children may be accessed in alphabetical order, such as the list available through the EPG from the satellite company Sky (Figure 3).

Figure 3. EPG from the satellite company Sky ©

Another way to find content via digital TV is by selecting a predefined category to choose children’s video on demand, such as movies and classic cartoons, available through the EPG from the cable company Virgin Media (Figure 4).

Figure 4. EPG from the cable company Virgin Media ©

The BBC iPlayer (www.bbc.co.uk/iplayer) allows users to browse for video content on-line alphabetically, by typing a keyword on a search box, or through exploration of content according to programme categories or the date of its broadcast (Figure 5).

Figure 5. EPG from the BBC iPLayer ©

The Sesame Street website (www.sesamestreet.org/videos) is target to preschoolers and also provides users different ways to search for video content. Keywords may be typed on a search box to allow search for specific content or the search and browsing may be done by subject, theme, character or songs (Figure 6).

Figure 6. EPG from the website Sesame Street ©

Thus, a functional requirement for an electronic programme guide is to allow users to search for specific videos as well as browse the content available in order to choose a video to watch.

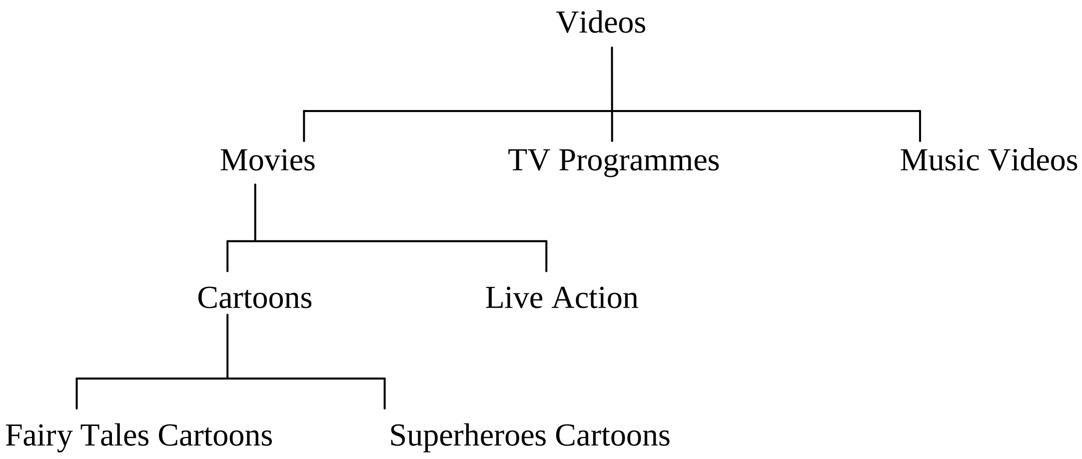
REA1 – Allow users to search and browse video content

Bakhshi (2007) recommends two complementary methods in the context of searching for video. These are the ‘bag of words’ approach, where viewers can search for content by entering keywords, and the ‘Subject/Genre/Channel’ approach where viewers can choose videos within a set of pre-defined categories. The first approach suggested would be hard to be implemented for pre-school children. Even with a speech interface such as that developed in the VISTA project (Carmichael et al., 2003). Children’s limited vocabulary and possibly idiosyncratic speech patterns would make the task complex for both system and user. Searching as well as browsing via pre-established categories is more viable and adequate for this age group.

REA2 – Provide searching and browsing via pre-established categories

In order to facilitate access and use of EPGs, like all information retrieval systems, classificatory techniques should be employed, this means that things should be grouped together according to common characteristics (Hunter, 2002).

According to Hunter (2002) the systems of classification may adopt a hierarchical approach, a faceted approach or a mixture of both. The hierarchical approach utilizes a series of groups or classes in successive subordination. In faceted classification concepts are analysed and grouped into facets giving users the ability to find items based on more than one dimension. The videos available on an EPG, for instance, could be divided hierarchically so that each entity (for example Fairy Tales Cartoons) falls into a sub-group of a larger group (Cartoons), which in turn forms part of an even larger group (Movies).



A faceted classification, on the other hand, lists within facets the terms that make up the subject. For example:

Video Types	Video Formats	Video Genres
Movies	Cartoons	Fairy Tales
TV Programmes	Live Action	Superheroes
Music Video		

In this case Fairy Tales Cartoons is not listed; it would have to be assembled by combining Cartoons and Fairy Tales.

The examples above were only partially developed to illustrate the difference between the two methods.

“No classification scheme will be suitable for all purposes and the choice, or design should be governed by factors such as the type of information system, the objectives of the system, and user requirement” (Hunter, 2002, p. 132). Research shows, however, that faceted classification reduces complexity and should be the basis of all methods of information retrieval (Broughton, 2006). In addition, Hutchinson (2005) indicates that flattened faced structures presented simultaneously are significantly faster, easier, more likeable, and preferred to a hierarchical interface for children to perform search tasks. Thus, a faceted approach should be adopted for the system of classification of the prototype application.

REA3 – The system of classification should adopt a faceted approach

In order to compile a faceted scheme first the relevant concepts have to be analysed and grouped into facets according to appropriate characteristics. These facets will then be treated as the EPG categories. Nevertheless, conceptual categories should not be merely characterized in terms of objective properties of category members. Human conceptual categories have properties determined by the nature of the people categorizing and have properties that are a result of imaginative processes (Lakoff, 1987). It was therefore found that children should be involved in the process to define categories for the EPG (see Section 4.4) in this way categories defined would be appropriate to the user group, making both searching and browsing tasks easier to accomplish.

2.3.2.3 Personalized Recommendation Systems

The main goal of an EPG, as described in the previous section, is to allow users to search and browse video content assisting them finding what to watch. Due to the increasing amount of content available, however, research has found a need to complement traditional EPGs with personalized recommendation systems (Xu and Araki, 2005). Hence, instead of having to search or browse through all videos available users are presented only appropriate content through intelligent interfaces, making the task easier to accomplish.

Studies on personalized recommendation systems offered some inspiration for the initial prototype design, but needed adaptation. The EPG-Board, for instance, is a social EPG that integrates a to-watch planner with a message board and a rating and tagging system (Iatrino and Modeo, 2007). The to-watch planner is composed by a reminder and a shared EPG. The reminder works like a planner in which the user may mark the programmes s/he wants to watch and the shared EPG displays the most interesting programmes for the user's buddy-list. The message board in its turn allows users to communicate in real time or in asynchronous assigning comments to programmes that may influence other users' decision on what to watch. The programme rating is a more explicit feedback in which users can say if they like the programme or not. Through the ratings the popularity of each programme is evaluated and incorporated into the shared EPG feature of the to-watch planner application. Finally, the content tagging application allows users to search content by browsing tags that were added by other users.

For the preschool audience, it would be difficult to implement a to-watch planner; to set reminders could be complex for this age group. It would also be hard to include a message board or a content tagging application aimed at pre literate users. The programme rating, however, could be feasible and useful. Three and four years old children could provide a feedback stating if they liked a programme or not and this information could be used by the system to recommend other programmes or to allow users to have easy access to the programmes they like.

REA4 – Allow children to rate video content indicating their preferences

The SenSee Framework combines the context with user profile to improve the multimedia consumption experience (Aroyo et al., 2007). Aroyo et al. (2007) highlight the importance of adding context including current time, geographical location and/or audience on top of content based filter for personalized search. The fact is that factors such as age and the time of the day may indicate different interests and this information can then be used to provide a context-aware personalised search and recommendation functionality. Thus, an EPG for young children should include only videos aimed at their age group. In addition, users' location and the time of the day should be taken into account. Content produced regionally will probably be more relevant to users than imported content, so it should be more easily accessible. The time of the day may also indicate which videos should be recommended by the EPG application, around children's bedtime, for instance, the recommendation should include calm relaxing videos to sooth children before their sleep.

REA5 – Only children's video content should be included in the application
REA6 – Regional videos should be more easily accessible than imported videos
REA 7 – Video recommendations should take into account the time of the day

AIMED is another personalized TV recommendation system based on user characteristics such as activities, interests, moods, experience and demographic information (Hsu et al., 2007). It makes programme suggestions using personal profile including demographic, lifestyle, and contextual information such as mood and viewing behaviour. In order to do so all first time users need to fill out a questionnaire which gathers demographic data, lifestyle tendencies and TV programme preferences when they register for the system. It would certainly be hard for preschooler to fill out such a questionnaire but it could probably be done by parents. The study reported in this thesis, however, did not involve parents, so it was decided that the prototype would focus only on children's ratings indicating their preferences (REA4).

2.3.2.4 Interface Concepts

Some requirements for system features were established in the previous sections. In this section some interface concepts will be discussed.

Currently the EPGs (Figures 3-6) have textual interfaces that are inappropriate for young children to search and browse videos on their own.

Hutchinson (2005) offered an interesting approach to a searching and browsing interface more appropriate to children's skills and preferences than traditional EPG interfaces. Hutchinson's research, along with the team from the University of Maryland, defined the current design of the International Children's Digital Library (Figure 7) (http://en.childrenslibrary.org). It provides more visual than textual information and is composed by a flattened faceted structure (REA3) presented simultaneously. Children may explore multiple, single-layer categories simultaneously that reduces the amount of navigation, removes abstract categories and facilitates searching and browsing.

Figure 7. International Children's Digital Library ©

The International Children's Digital Library, however, is for book searches aimed at a wider age group while the EPG prototype would be for video searches only for three and four year olds. For this reason several categories could be eliminated such as the first three categories on the left concerning children's ages: three to five, six to nine and ten to thirteen. In addition, the categories on the top could also be excluded because they regard the colour of book covers. Children could probably choose DVDs by their cover colour, but for television content such concept is inappropriate. Thus, for the prototype under development it would not be necessary to include so many categories divided into two different pages. As a result, it was decided that the number of categories could be decreased to fit in only one page (RL17).

As previously mentioned children are involved later in the research to assist on the establishment of categories. At this stage, however, categories were pre-established according to categories from the International Children's Digital Library as well as children's channels websites and children's films.

The International Children's Digital Library presents 44 categories spread over two pages. Some categories were found appropriate for video content: make believe, real animal characters, fairy tales and folk tales, series, songs and fantasy.

A few children's channels websites were very briefly analysed and also provided some inspiration to establish categories. CBeebies website (http://www.bbc.co.uk/cbeebies), for instance, includes categories such as animals and nature, around the world, make and do, music and songs and stories that would be useful for searching and browsing video content. Disney website (http://home.disney.co.uk) displays the categories movies, TV and music. CITV, Cartoon Network, Nick Jr. and Boomerang provided similar categories. CITV (http://www.citv.co.uk) has a category named create it equivalent to CBeebies' and Nick Jr.'s (http://www.nickjr.co.uk) category make and do. Nick Jr. also presents the category shows, while Boomerang (http://www.boomerangtv.co.uk), Cartoon

Network (http://www.cartoonnetwork.co.uk) and Jetix (http://disney.go.com/disneyxd) include the category videos, similar to Disney's TV category.

Traditional EPGs include in the category kids the subcategories movies, classics cartoons, along with children's channels for children to choose content according to the channel it was originally broadcast.

Based on the existing applications mentioned above and some of the requirements gathered thus far, fourteen categories were pre-established for the prototype application: movies, cartoons, music, make and do, animals and nature, fairy tales, superheroes, around the world, TV shows, children's favourites. And the channels aimed at preschoolers broadcast in the UK: CBeebies, Playhouse Disney, Nick Jr. and Cartoonito.

Having decided the categories to be included the following step was to decide the way to present them.

According to Shneiderman (1998) the central principle for visual information seeking is: "overview first, zoom and filter, then details on demand" (Shneiderman, 1998, p.523). The user should first be given an overview of the entire collection, after that s/he should be able to zoom in on items of interest, filter out uninteresting items and finally select an item or group to get details when needed.

In order to give an overview of the videos included in the electronic programme guide it was decided to present them in a dynamic 3D wheel such as the one used in Disney's website to access different characters' mini-sites (Figure 8) which children are already familiar with from the websites for their age group. As well as being familiar to children, this form of presentation also enables more content to be displayed than a static presentation such as the one used in International Children's Digital Library (Figure 7).

Figure 8. 3D Wheel from Disney website ©

The 3D wheel would provide an overview of the content available and the screenshots presented on the wheel would be used to access the video content. The wheel is for this reason the most important item in the interface; it presents and gives access to content. Thus, it was decided to display the 3D wheel in the middle of the page (RL26).

The categories could be used to filter the content to be displayed on the wheel. The categories were organized on the page so that those equivalents to video types and formats would be presented on the first row, then on the second row would be displayed the video genres and on the third and last row the children's channels categories.

It was found complicated to provide preschoolers video details for their selected choices, most details on EPGs and information retrieval systems are provided in textual format, completely inappropriate for young children. For this reason, after given an overview of video content available, children would be able to filter the collection and then select a video to watch instead of obtaining details of their selection before choosing. In case they were not happy with the choice made they could then go back to the main menu and select another option.

The way the navigation would work was then established through the review of existing applications and requirements gathered. The look of the interface, however, including colours, icons and final set of categories was to be decided in the following stages of the research along with children in order to reflect their preferences.

2.3.2.5 Conclusions

From the review of existing applications seven requirements were established and are listed in the Appendix A. In addition to the requirements, existing applications also offered inspiration for the initial interface design presented in Chapter 4 (Section 4.2).

2.4 Conclusions

According to the background literature presented in this chapter it may be concluded that the landscape of children's television has changed; children today are offered a vast number of choices via the television and computer. In this scenario it is crucial for an interactive television application to enable young children to choose appropriate content on their own using different input devices.

Due to the lack of specific literature on interactive television applications for young children, the initial process of development of the prototype application was based on research on child development, existent guidelines for interaction design and interactive television, and analysis of existing applications. Thus, the process of review of relevant research was useful to emphasise the need for the work described in this thesis, and also to provide a list of initial requirements for the prototype application (Appendix A).

Chapter 3. Research Design

3.1 Introduction

This chapter contains the main issues related to the research design adopted in the work described in this thesis. Section 3.2 contains a restatement of the research aims and a brief overview of the research approach. In Section 3.3 the research design is explained including details about the development of the prototype and children's involvement in the process. Section 3.4 contains information about the implementation of the research design considering, amongst other things, the selection of technology and participants. The conclusion in Section 3.5 summarizes the chapter and provides an overview of the rest of the thesis.

3.2 Research Approach

As stated in the previous sections, the aims of this study were to analyse how preschoolers interact with iTV applications in order to develop a reference framework, infer design principles and refine methods of design and evaluation; in order to achieve these aims an empirical approach is used.

The emphasis of the research is to discover concepts and relationships in raw data and then organize these into an explanatory framework of children's interactions, derive principles and refine methods. Most of the analysis is qualitative and interpretative, although one of the design activities (closed card sorting, see Chapter 4) included a quantitative approach.

This empirical research, based on observation and experimentation, was inspired by work in educational design research (Akker et al., 2006).

Reeves (2006) reviews design research from a technology perspective and states the primary advantages of a design research approach as follows:

- The identification of real problems
- The creation of prototype solutions based on existing design principles
- The testing and refinement of both the prototype solutions and the design principles until satisfactory outcomes are reached

Reeves argues that instead of conducting predictive research in isolation from practice, in which translating findings into practical solutions is complex, it is better to undertake design research integrating the development of optimal solutions to problems in context with the identification of reusable design principles.

Educational design research is defined as "a research approach suitable to address complex problems in educational practice for which no clear guidelines for solutions are available. Educational design research is perceived as the systematic study of designing, developing and evaluating educational interventions" (Plomp, 2007, p.9).

It was found that this approach was suitable for the research presented in this thesis because the purpose of the study was to design, develop and evaluate an interactive TV application aimed at young children, for which there are no clear guidelines available, integrating solutions to problems found during evaluation sessions with the identification of design principles.

Hence, informed by prior research on interactive television and interaction design for children, a prototype was developed by carefully studying successive versions of the prototype in collaboration with children, and during the entire process children's feedback was analysed in order to produce design principles. An iterative design research process was followed and encompassed analysis, design, evaluation and revision activities iterated until a satisfying balance between ideals and realization was achieved (Plomp, 2007, p.13).

The research was, however, only inspired by design research instead of based on the approach. The main difference between the design research approach and the research conducted is that the research presented in this thesis was not focused on educational aspects of technology.

The research results are presented in a framework that indicates factors that influence children’s interaction with an iTV application. Findings from the design and evaluation processes were used to generate design principles for interactive television for young children. Techniques for design and evaluation sessions were refined throughout the process.

3.3 Research Design

This section outlines the research design used in this project. The model presented by Reeves (2006) (Figure 9), although not followed rigorously, serves to demonstrate the process of data gathering and analysis, prototype development and production of design principles that was used in the research work.

According to Reeves, practitioners and researchers should (1) collaborate in the identification and analysis of problems, (2) create prototype solutions based on existing principles, (3) test and refine both the prototype solutions and the design principles until satisfactory outcomes are reached (4).

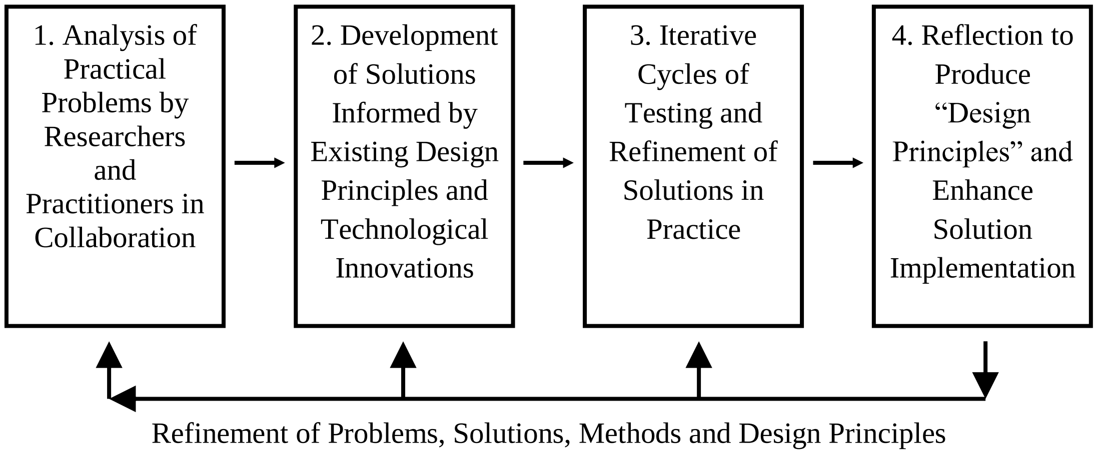


Figure 9. Reeves' (2006) illustration of design research approach

As in Reeve's (2006) approach the research work reported in this thesis started with an analysis of the problem, how iTV applications can meet young children's needs, abilities and preferences. This was done, however, through the literature followed by design activities instead of through analysis by researchers and practitioners in collaboration, because the study was not focused on educational interventions. During the design activities, requirements were identified and solutions for the design of the prototype proposed. The process of design activities that led to additional requirements and prototype improvements was iterative and continued until most of the prototype's characteristics had been analysed and revised. The evaluation stage, on the other hand, was not composed of iterative cycles. Differing from Reeve's approach, the iterations in this study were carried out before the testing. Following the evaluation sessions there was then reflection to produce the framework, design principles and refine methods for design and evaluation with young children; these were also informed by previous stages of the research, such as the literature review and design activities. The diagram below (Figure 10) illustrates the process described.

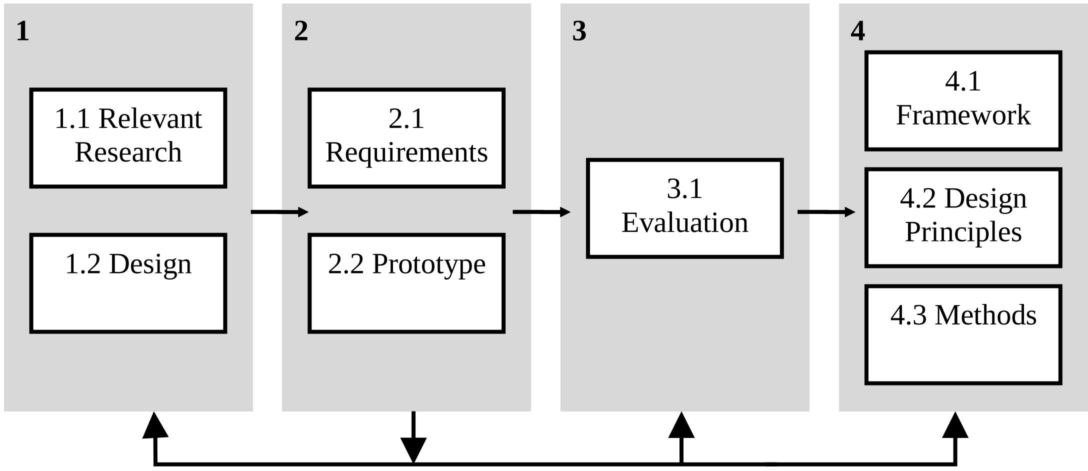


Figure 10. Illustration of the research design approach taken

3.3.1 The Development Process

In the previous section an overview of the research design approach taken was presented. In this subsection the development process of the prototype application is explained in more detail.

There are several different ways to develop interactive applications; in the research presented in this thesis a user centred design approach is used.

A model traditionally used in software development projects is the linear and sequential *waterfall* model (Somerville, 1995). This approach is composed of different stages: requirements analysis, design, code, test and maintenance. Once a stage is completed, the development proceeds to the following stage and the preceding steps are not reviewed at any point. The requirements, for instance, are analysed during the initial phase and remain unchangeable throughout the development. There is no iteration, so they are not re-evaluated in subsequent stages. It is widely accepted that such an inflexible development process is inappropriate for designing interactive systems (Markopoulos et al., 2008, pp. 36-37). For this reason, most versions of the *waterfall* model currently used incorporate some level of iteration (Preece et al., 2002, p.188). However, the model still does not offer the opportunity for evaluation involving users. It is therefore inadequate for the development process of the work described in this thesis.

A user centred approach demands not only an iterative design but also user participation. It starts when the need is identified for a human centred project and involves an iteration of the following interdependent stages: specification of context of use, specification of user requirements, production of design solutions, and evaluation of the design against the requirements; these stages are then iterated until the requirements are met (ISO-13407, 1999). This human centred design process is equivalent to the process of interaction design described by Preece et al. (2002), which involves four basic activities:

- Identify needs and establish requirements
- Develop designs that meet the requirements
- Build interactive versions of the design
- Evaluate what is built throughout the process

The authors highlight three characteristics that should be part of the design process: users should be involved throughout the development; usability and user experience goals should be identified at the beginning of the project; and there should be iteration throughout the four activities.

The development process followed in this project (Figure 11) was based on the iterative user centred approach described above, in which the iterations were used not only to enhance the prototype but also to contribute to the composition of the framework of children’s interactions as well as design principles.

Relevant research (1.1) informed the requirements for the prototype application and also contributed with guidelines, techniques and ideas to structuring design activities and evaluation sessions. The design decisions were based on requirements (2.1) and gradually composed the prototype (2.2) through iterations in which design activities were conducted to inform further requirements and design decisions that were then fed back into the prototype. Following the iterations, as soon as the main prototype characteristics were defined, the final version of the prototype was evaluated (3.1).

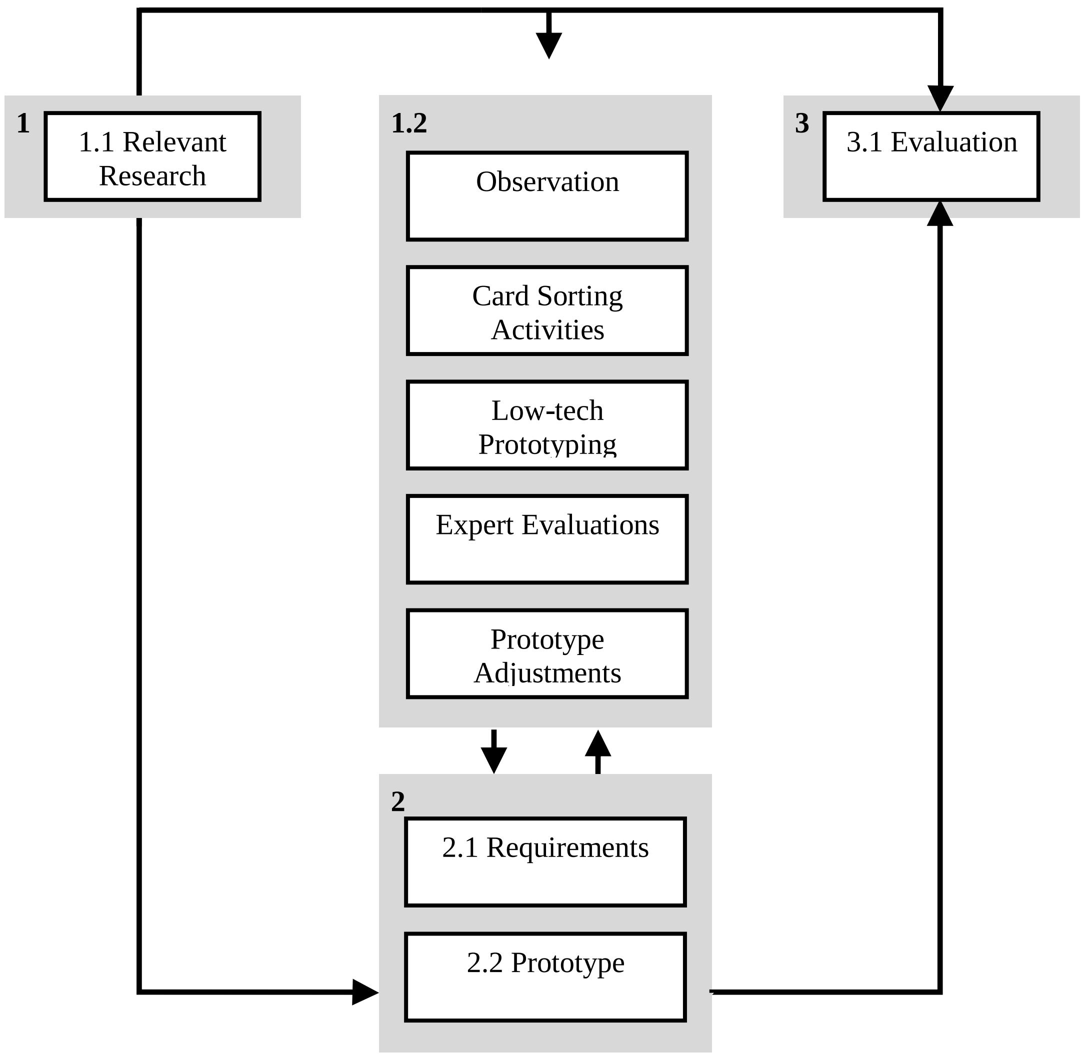


Figure 11. Stages of the iterative development process undertaken

3.3.2 Children's Involvement in the Process

As mentioned in the previous section, in the process of interaction design, user involvement is crucial. Children may, however, participate in different stages and play distinctive roles during the process. Druin (2002) defines four different roles children can play in technology design: user, tester, informant and design partner. As users, children use *existing technology* so that future technologies can be enhanced based on the analysis of their experience. In the role of testers, the minimum involvement required for a user-centred development (Markopoulos et al., 2008, p.45), children test *prototypes* that are later improved according to the results from the testing sessions. As informants, children are asked for input and contribute to the *design* at several stages. In the role of design partners, equivalent to *participatory design*, children participate throughout the entire process in which they are considered equal stakeholders, so design decisions are negotiated between children and adults (Druin, 2002).

In the work described in this thesis, the informant design approach was chosen, it resides between user centred design and full participatory design (Scaife et al., 1997). Taking a participatory design approach in which children are design partners would be time consuming, require more researchers, and probably would veer the focus from developing a framework as well as design principles towards specific design solutions. On the other hand, children participating as users or mere testers would not elucidate as much detail of their interaction with the media. When involved earlier in the research children are more likely to give suggestions and talk about different ways of interaction than during testing sessions with high-tech prototypes (Scaife and Rogers, 1999). Thus, the informant design approach was found the most appropriate both to contribute to the prototype enhancement and development of design principles and to assist in the constitution of the framework. As a result, children contributed during several design activities and then evaluated the prototype.

3.4 Implementation of the Research Design

In this section we describe some of the practical decisions made with respect to the research and the impact they had on the validity of the results.

3.4.1 Selection of Technology

As previously stated, interactive television is defined as programmes, applications and services that the user is able to interact with, altering the audiovisual content being displayed on the screen. This interaction could take place using different devices such as an interactive toy, a keyboard, a remote control, a mouse, or via a touch screen.

Initially, the idea was to develop a system in which the user could interact with the television content through a plush animal. During the early stages of the research, a prototype was developed of a plush baby bear for which the user would be responsible and to which he would teach numbers, colours and letters by watching different types of content on the television accompanied by the baby bear. The "father bear" would be an interactive television application and would ask the user to perform tasks to ensure baby bear's development. As the user watched the educational content, with the baby bear, the plush animal would "learn" the content and "develop". This meant the child would be able to increase the number of songs baby bear could sing and numbers it could count to by watching television. The type of content presented on the screen would alter the content available through touch sensors on baby bear, and the interactivity with the baby bear would alter what father bear presented on the screen.

Considering this technological approach two main concerns were raised. The first concern was about the design of the study involving interaction via this plush animal; it would be problematic to evaluate the user experience in a single session with each child. It would be necessary to ask participants to interact with the toy for a period of time so that the concept of the “learning” plush animal could be concretized. This would involve more participant time and would probably result in more interesting data if children were not directly observed, but were video recorded continuously while playing with the toy, with and without the television, which would raise several ethical issues. The second concern raised was about the usefulness of the results obtained for a framework of children’s interactions with interactive television applications. The use of the plush animal could reveal issues attributable to the interaction via touch sensors, for instance, that would be very limited to interactions of this type and could offer no enlightenment for interactions via more conventional devices such as the remote control.

Comparing the mouse and touch screen for children’s use of computers Lu and Frye (1992) stated that touch screen is a more effective input device for preschoolers (Lu and Frye, 1992). Romeo et al. (2003), however, reported that most young children in their study preferred using the mouse instead of the touch screen as an input device; the authors affirmed that the familiarity and proficiency, of children as young as three years old, with both mouse and keyboard affected children’s use of the touch screen (Romeo et al., 2003). The touch screen was also considered as an input device in the study reported in this thesis; it was, however, discarded because it is not widespread in television and computer usage within a home environment. For similar reason, Wiimotes and joysticks were not included.

Currently, most of the interaction with audiovisual screen content happens via the remote control (on digital television and set top box), or through the mouse and/or keyboard (on the computer). In order to obtain data that could contribute to a reference framework and to design guidelines for these common interactive applications for children, it was decided to select more conventional devices for this research. In order to simulate the television and computer environments, participants were asked to interact with the remote control and mouse or keyboard.

For the evaluation of the prototype it was decided to use a computer instead of a set top box. As mentioned in Chapter 2 (Section 2.2.2), screens resolution and viewer/user behaviour interacting with the television and computer are similar. Thus, the computer was chosen for practical reasons, since it would be easier to carry a laptop to the nurseries than a set top box together with a TV set, and because it offered the flexibility to conduct the evaluation with both input devices; an infrared PC remote control and the mouse.

3.4.2 Rationale for Choice of Application

The technology choice was explained in the previous section. In this section the selection of the application to be tested using the remote control and mouse is justified.

In order to analyse children's interactions with iTV it was decided to develop a prototype for which children could contribute to the design and evaluation.

A number of applications were considered for the prototype, from interactive narratives to drawing tools. It was eventually decided to develop an electronic programme guide (EPG) prototype, since it is a vital component of iTV that enables viewers to find their way in an ever-increasing landscape of audiovisual content.

The EPG is a guide to television programmes displayed on the screen allowing viewers to search or browse and then select programmes to watch from the options presented. The design of EPG's has been addressed by a number of studies, focussing variously on efficient EPG design (Daly-Jones, 2003, Eronen and Vuorimaa, 2000) and personalized recommendation systems (Aroyo et al., 2007, Hsu et al., 2007, Iatrino and Modeo, 2007), yet this research tends to assume a homogeneous adult audience.

Television viewing is part of preschoolers' daily routine and is often an activity performed unaccompanied. Since children's requirements are different than adults', they would benefit from a specialised EPG. Thus, besides contributing to the framework of children's interactions with iTV, the choice of EPG for the prototype application could also offer insight for future designs of electronic programme guides aimed at young children.

3.4.3 Sample Sizes

Different sample sizes were used in different stages throughout this thesis. During most of the design phases in which children contributed as informants, a small sample was found adequate, since the activities, typical of HCI practice, were meant mostly to inform the design of the prototype.

Well known HCI methods and tools, described in the following sections, were used during the observations, low-tech prototype session and prototype adjustments with children. However, the card sorting activities were developed based on examples from the field of psychology. There were several concerns about adapting the study to contribute to the design process. For instance, it was uncertain if children would be able to categorize audiovisual content, if they would understand that screenshots were representing videos and if the activity would actually work. As a result, in order to refine and validate the card sorting task, a larger sample was found necessary. This sample size was defined with the assistance of a statistician using the tables for logistic regression (Hsieh, 1989).

During the evaluation phase two options were considered for conducting the study with a working prototype: (1) a large sample from several nurseries and data captured during brief usability testing sessions to assess children's performance while interacting. While this would result in high external validity, it would not permit a rich description for the analysis in detail of each child's interactions. (2) A small sample in which children's interactions could be analysed in depth. This approach affects the external validity but was found more appropriate to provide the details needed to describe how the interaction occurred and assist in the interpretation of a phenomenon not yet well understood.

This research results in a framework, developed from the study of a few children. It does not have the explanatory power of a more general and larger theory. The sample, however, was made as systematic and widespread as possible, so more conditions and variations could be discovered and built into the framework to enhance its explanatory power. Furthermore, a rich and full description of the context in which the research was conducted is included in the following sections of the thesis to allow reproducibility and transferability.

3.4.4 Selection of Participants

The participants who contributed during most of the design process were children from the One World Nursery at the University of Brighton (UK). Conveniently located on the University campus, the nursery offered access to children from diverse backgrounds.

To refine and validate the card sorting activities, sessions were conducted in an additional five nurseries also in Brighton. In order to select these nurseries first letters explaining the study were sent to all 46 nurseries in Brighton. One week later the nurseries were contacted by phone. Some of them declined to participate and in others managers could not be contacted at on the first attempt. Nevertheless, five of them agreed to take part in the study and provided the sample needed.

For the evaluation phase, it was decided to involve children who had not participated during the design process, in order to verify if the design decisions assisted by those participants, such as categories and icons, were significant to other children as well.

Research in interaction design for children highlighted the increasing digital divide between children in developed and developing countries as well as across socioeconomic lines within the same country, and indicated “a need to broaden target populations across social, economic, and cultural lines” (Hourcade, 2008).With the aim of broadening the population and selecting as broad a sample as possible, in order to discover conditions and variations and enhance the explanatory power of the research, it was decided to conduct the evaluation in two different countries, in the United Kingdom and in Brazil.

The initial idea was to conduct the evaluation in a small town in the Brazilian countryside with low socioeconomic status, in which most children would not have access to digital television or computers. The aim here was to assess if their familiarity with analogue television would be sufficient to provide them a good user experience while interacting with the interactive TV prototype. This would involve, however, teaching young children how to use the mouse, which would be time-consuming and probably not practicable in a single session. Another issue raised with this approach was that it would be difficult to find a similar condition in the United Kingdom, where most nurseries are equipped with computers and the majority of the population has access to digital television (Ofcom, 2009) and/or computers as well as internet (Statistics, 2009). It would be interesting to have a sample as wide as possible, but several cultural differences could emerge while studying Brazilian and British children with similar socioeconomic status and access to technology. Extreme variations within the participants could present problems such as the inability to use the mouse and would offer no basis for a framework.

For this reason, São Paulo was the city chosen to conduct the study in Brazil. This is a large city that offers diversity and is the Brazilian region in which more people in comparison with other parts of the country have access to computers (IBGE, 2009) and digital television (DTV, 2010). In order to include children from different socioeconomic backgrounds in the evaluation it was decided to involve children from two nurseries, one private in an affluent neighbourhood (charges monthly fees), and one public in a poorer neighbourhood (free). The number of children from each nursery, however, was not equally split; the majority of children were from the public nursery, due to the number of signed consent forms returned.

The city chosen for the evaluation study in the United Kingdom was London, because it is a large city such as São Paulo, which offers diversity. Since most of participants in Brazil were from a low income neighbourhood it was decided to conduct the study in a low income neighbourhood in London as well. To include some variance in children's socioeconomic background the evaluation was conducted in a voluntary and in a private nursery. Both nurseries charge average weekly fees but in the UK three and four year old children are entitled to free early years education for twelve and a half hours per week in private, public, voluntary and independent nurseries as well as financial help with the costs of extra childcare.

There were ethical concerns on socioeconomic lines therefore we did not try to make a rigorous segmentation.

To select the nurseries to participate in the evaluation phase, the same procedure used during the card sorting activity would be followed, i.e. letters and then phone calls. It was found, however, that due to the postage time to Brazil, e-mails explaining the research would be faster than and as effective as letters. Thus, e-mails were sent to all public schools in lower working class neighbourhoods and all private schools in three randomly chosen middle class neighbourhoods in São Paulo. In London e-mails were sent to all nurseries in one lower working class neighbourhood. One week later the nurseries were contacted by phone. Those who agreed to participate were sent the consent forms to be distributed and signed by parents.

During each stage of the research, the inclusion/exclusion criteria used were children's age group and a signed consent form. A written consent form was sent to parents and carers explaining the study and asking for permission for their children to take part in the research. Children were also told about the activities and asked if they would like to participate. The participants were told that their participation was completely voluntary and they could withdraw the consent to participate at any time during the process.

As a result, children who took part in the study were three and four year olds whose parents had signed the consent form allowing them to participate and who wanted to take part after the research was briefly explained to them.

3.4.5 Research Practice

Before the data collection started, the research project was approved by the University Ethics Committee. To protect children's privacy, all data collected was anonymised and confidentiality maintained. Children were referred to by their first name during the activities, but on the data collected their names were replaced by a participant number and the data analysis included only these numbers. During some stages of the research children were videoed and the video files along with data that could identify children were stored in password protected computer. In these cases the parental consent form included an option in which they could give permission or not for their child's images being used in publications.

Every stage of the research was carried out in nursery environments. Some phases, such as the observation, were conducted in the main nursery room in which the children being observed were surrounded by other children. Other stages such as the card sorting activities were carried out in a separate room or in a corner of the nursery setting. The evaluation sessions were always conducted in separate rooms. On some occasions the researcher had to re-locate as rooms were needed for other nursery activities. For children to feel more comfortable during the study, when the activities were conducted in separate rooms the researcher was always accompanied by a member of the nursery staff.

In order to be inclusive and provide the opportunity to all children who wanted to participate to get involved, non-participants, those outside the age group or whose parents had not signed the consent form, were offered the chance to test the prototype and conduct the card sorting activities at the end of the experiments without any data being collected. One child with special needs participated in the research, but he was six years old and for this reason excluded from the data analysis.

3.5 Conclusions

This chapter described the research approach and design, inspired by educational design research, adopted in this thesis. The research design and iterative process of development of the prototype were sketched. Children's participation in the process of design and evaluation was explained and details about the implementation of the research were justified.

The remainder of the thesis falls into four sections. The design activities are described in Chapter 4. Chapter 5 presents the evaluation methods and results. In Chapter 6 findings from the two previous chapters are discussed. It includes the framework of factors that affect children's interactions, along with design principles and a review of refined techniques for design and evaluation studies with young children. The conclusion to the thesis is then presented in Chapter 7 with a summary of the main findings and ideas for future work.

Chapter 4. Design Activities

4.1 Introduction

Throughout this chapter the design activities that contributed to the development of the prototype application are described and discussed, design choices are explained, requirements derived from the studies are identified and design decisions proposed. These requirements, along with design solutions will be then verified during evaluation sessions (Chapter 5) to contribute to the framework (Chapter 6) and to be translated into design principles (also presented in Chapter 6).

The chapter is structured in a way that follows the process of development of the prototype. In Section 4.2 the initial idea for the prototype, based on the research carried out up to this point, is illustrated in a wireframe. Section 4.3 describes an exploratory study in which children were observed in the nursery setting. Section 4.4 includes novel card sorting studies used to define the categories for the prototype. In Section 4.5 low-tech prototype sessions, carried out to identify appropriate icons to be used in the interface, are described. A high-fidelity version of the prototype is presented in Section 4.6. Section 4.7 presents expert evaluations focused on improvements for the design of the prototype. In Section 4.8 prototype adjustments sessions are explained. Finally, Section 4.9 includes a brief summary of the chapter.

4.2 Prototype Wireframe

Wireframes are usually sketches of important screens in the application, software or webpage. The images contain crude representations of interface elements that appear on each screen along with explanatory text describing the screens and indicating what happens when a particular control is activated (Silver, 2005).

At this stage in the research a wireframe of the prototype application was developed with the requirements gathered thus far. The main aim was to use it as a low-fidelity prototype, simple and quickly produced in order to illustrate the initial ideas for the interface and most importantly to support exploration of alternative designs (Preece et al., 2002).

Figure 12 illustrates the main prototype screen. The round buttons display the EPG categories. It is a flattened hierarchy (RL25), with no sub-menus (RL8) and all the content is displayed in a single page (RL17). The user may click or press the OK button on the remote on a category to select it (RL14). The selected category is highlighted (RL35) and the content of that category displayed in the 3D wheel. The screenshots are represented by the rectangular forms and move around the wheel. The user may click or press the OK button on the remote on the screenshot representing the video s/he would like to watch. The arrow buttons next to the wheel may be used to speed the wheel movement in both directions so that the navigation can be flexible and efficient (RL3, RL29).

In addition, the prototype includes a help feature (RL33) in which users may access information about the system and how the navigation works.

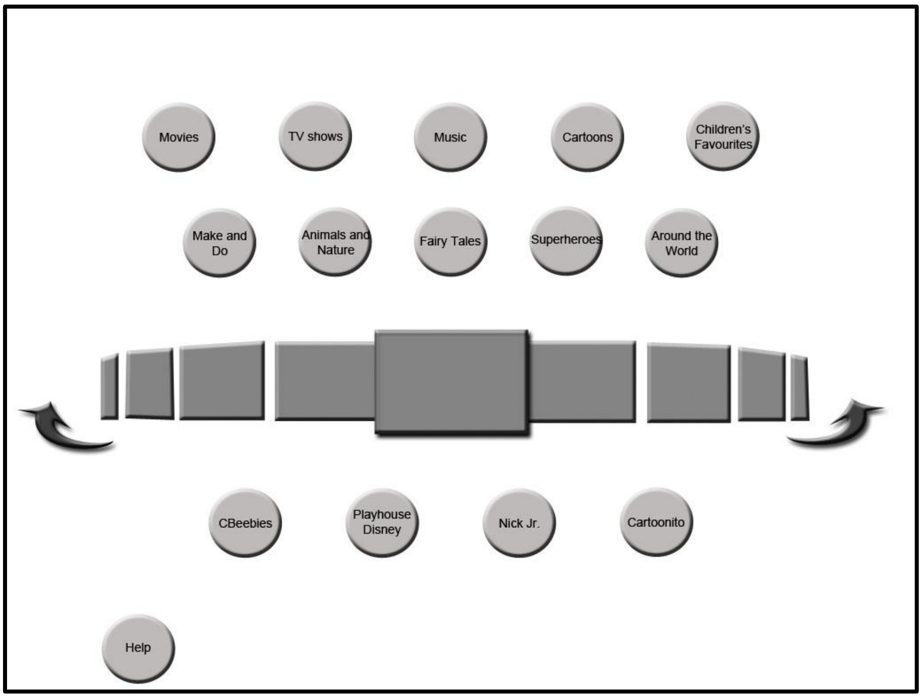


Figure 12. Prototype wireframe containing the EPG main menu

As soon as the user clicks or presses OK using the remote on the screenshot the video is loaded in full screen mode. Figure 13 illustrates the prototype screen in which the user is watching a video. The screen contains a button to return to the EPG main menu, another button to access the help section and an additional button to rate the video being watched (REA4).

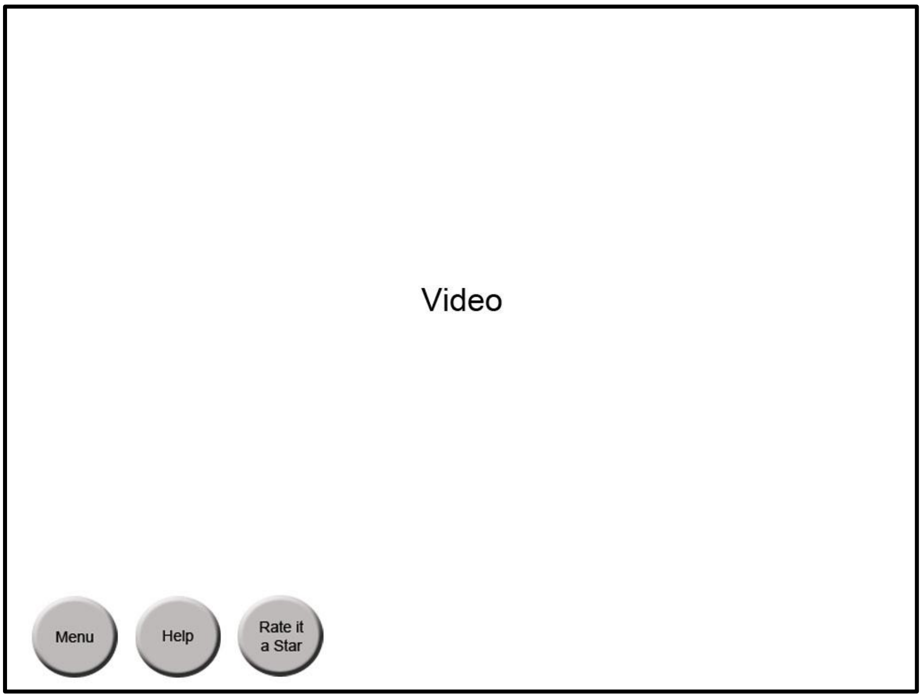


Figure 13. Prototype wireframe screen in which the user is watching a video

The details, colours, final categories of the prototype and icons will be decided in the next stages of the prototype development.

4.3 Observation

4.3.1 Introduction

Observation involves spending some time with users to get a richer view of their behaviour from which requirements can evolve. Observations may be carried out at different stages of the project development. Early in design process it is useful to assist designers with an understanding of users' needs. Observations may be done in a controlled environment or in the field and observers can be outsiders or insiders (Preece et al., 2002).

The objectives of observing children in the study reported in this section were: familiarization with the age group, an opportunity to know the children and their behaviour, for them to feel more comfortable with the researcher during the following stages of the study and also for some issues to be clarified, fed back to the prototype and assist on the analysis of children's interactions such as: do young children search and choose activities, books and games? Do they rate the activities/experiences afterwards? Do they suggest it to their friends? Do they need instructions? How do they look for instructions? Who provide these instructions and how?

To achieve these aims it was necessary to observe children on their natural environment, given that they would probably not behave as naturally in a lab. Analysing children at their homes would imply on several practical and ethical issues, for this reason it was chosen to observe children in the nursery.

It would certainly be complex to ask preschoolers to answer the questions presented above to clarify issues for the prototype application. Thus, in order to respond those questions and gain a richer view of children and their behaviour it was decided to take an outsider approach and engage in conversation only if it was initiated by children.

Children were then observed in the nursery setting and an outsider observer approach was taken during observations.

4.3.2 Method

4.3.2.1 Participants

Six children, three girls and three boys, between three and four years old, were observed in the nursery. This activity was conducted in the One World Nursery at the

University in Brighton (UK) and the inclusion criteria were children's age and a signed consent form. Children from this nursery contributed at several stages of the research, so the letters and consent forms sent to be signed by their parents included several activities (the parental information letters and consent forms are attached on the Appendices B and C).

4.3.2.2 Procedure

Children were observed during two weeks in three hours sessions twice a week; twelve hours in total were spent observing preschoolers. The data was collected in form of field notes, to be coded and categorized, so actions and procedures could then be defined and inform the framework and requirements to the prototype.

4.3.3 Results

The field notes were composed by descriptions of children's actions in the nursery setting and the descriptive details to be recorded and analysed were chosen based on what was thought to be important to the framework and prototype application, including situations in which children search and choose, recommend and ask for help. These descriptions were then organized into categories according to their properties and related to each other to form an explanatory scheme (Figure 14) integrating the concepts.

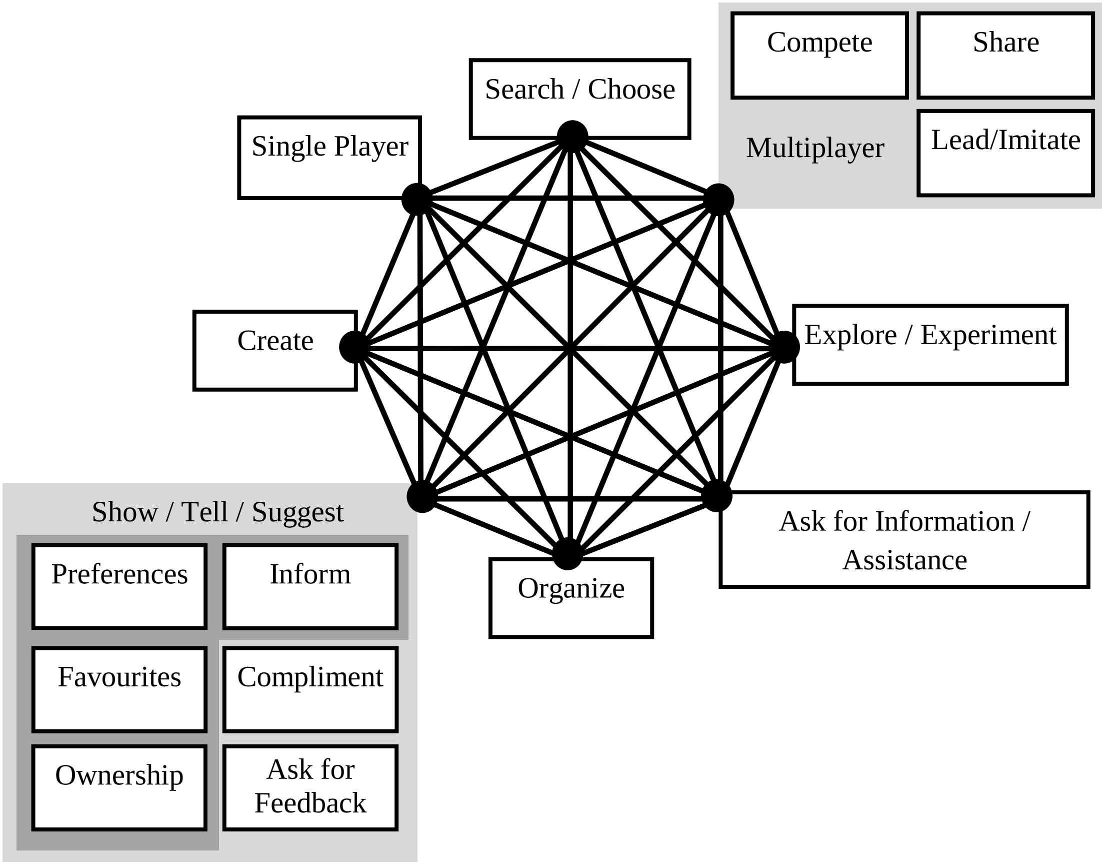


Figure 14. Children's actions and behaviours related to prototype features

The diagram above incorporates the structure children are given during most activities at the nursery. First they are provided with options to choose. Outside it includes different toys and objects such as hoppers and ped-a-rollers, and places like the sandbox and the garden, while inside they can draw, colour or paint; using a range of brushes and materials, play with clay, sand and water and also dress up and role play. As soon as children choose their activities they play it alone or with friends, in case they are playing with other children it may involve sharing (materials or experiences), competing, leading or imitating. During play, alone or with friends, children explore and experiment. They engage in physical and sensory experiences such as pouring water through funnels and shaking bells in different ways and directions. Children also create drawings, paintings, sculptures, and role play. During the activities children generally communicate with other children and adults. They show, for instance, the helicopter passing by to friends and members of the staff; tell "I have too much on my tummy" (participant 1 to a staff member) and suggest "don't go down there" (participant 5 to a friend). Preschoolers also compliment their friends, "Oh, that's clever" (participant 2 tells his friend); and indirectly ask for feedback showing the work they've created to adults or informing of accomplishments "I ate all my lunch" (participant 5 to a staff member). Children also talk about what they like, "I like decorating trees" (participant 3), their favourite things such as colours, food and animals and their ownership, "Oh these are my crumbs" (participant 1 tells her friend during a role play in which they are making dinner). Preschoolers generally ask adults and other children for information about things they are not sure what it is, "Hey what's that piece?" (Participant 1 asks her friend referring to a toy on the sand), or where it belongs "Where all the kettles go?" (Participant 3 asks a member of staff). They also ask friends for assistance to expel a younger child from their play, for instance, and ask adults for help on putting their aprons on and opening containers with their lunches. At the end of the activities there is usually a tidy up time in which children organize the materials they were playing with.

The eight categories included in the scheme are all interconnected; children usually start to play by searching and choosing an activity and end the process organizing the setting. There are however several activities that involve searching, organizing or both as part of the play, such as puzzles, for this reason the scheme cannot be represented in a hierarchal manner. Play modes are also connected and change constantly, a child may start to play alone then a group of friends join in the fun, then leave, and the child continues to play alone. The type of play, exploring/experimenting, crating/role play is not only connected but can also happen simultaneously. Two girls making cakes using plastic containers and sand are role playing, creating and experimenting during the same activity. The beginning of this activity could be then represented on the diagram 2 with a triangle uniting the three vertices "Multiplayer", "Experimenting / Exploring" and "Creating / Role Playing" that later would also incorporate “Show / Tell / Suggest”, “let’s make some dinner” (participant 1 suggested), and subsequently “Ask for Information / Assistance” when one of the girls asks the other to tell the younger child approaching she should leave.

The categories representing actions and behaviours in the scheme are then interconnected, dynamic and can happen concurrently; it would be ideal to provide a type of activity with the system that would fill out the entire area of the octagon, but this was not possible. The scheme certainly offered insight to prototype’s features and assisted on design decisions, but some concepts could not be incorporated.

The system presented the possibility for the user to search, browse and choose videos to watch. There was, however, a concern that there were too many categories and elements on the screen and children would be overwhelmed. Research suggests that the number of menu items presented should be between three and eight because on average a human short-term memory can register five to seven chunks of information, more than that could overwhelm users (Kunert, 2009). Gawlinski (2003) recommends there should be less than seven elements on screen at any time. The first version of the prototype (Figure 12), however, was composed by more than 20 items. Nevertheless, observations at the nursery indicated that children search and choose activities among an extremely large amount of options, sometimes they are limited to the room they can play, but there are still more than twenty options of toys, materials and activities available.

In the nursery art room, for instance, where children can choose materials to paint and draw, there are two shelving units (Figure 15) containing different types of materials for them to use. From paint to dried pasta each container has a different material. The two shelves on the bottom of each shelving unit are easily accessible to three and four year olds. Anything that children can reach they can use, the materials they cannot reach are usually provided by members of the staff.

Figure 15. Shelves with a range of materials for children to use in the art room

From the figure above it may be noticed that even if only the materials children can reach are considered the number of possible items they have available to use on their art work exceeds 20. For this reason, it was decided that the categories and items in the interface did not need to be limited by a specific number, they should instead be plenty and allow children to explore it.

RO1 – A large number of options should be provided

RO2 – Provide room for exploration and experimentation

The use of operators like 'and', 'or' and 'not' in search systems is very common and this type of searching is known as Boolean searching (Hunter, 2002). It was decided that only one operator, the 'and', would be used in the prototype because it would be hard to indicate to young children both conjunction and disjunction. In addition, the operators 'or' and 'not' would probably increase the number of videos resulted instead of filtering the results. Thus, the prototype application would then provide a possibility for conjunctive Boolean search in which users could select two or more categories to narrow down the number of videos presented, filtering their selection to choose.

Hutchinson (2005) states that younger children do not always understand they are creating conjunction Boolean searches when selecting more than one category on the interface. Nevertheless, it was decided to provide opportunity for Boolean searches to cater to more experienced users providing flexibility for the system (RL3) and also to provide a way for children to explore the system. The exploration feature of the prototype was then considered to be the different categories with the possibility for conjunctive Boolean searches.

RO3 – Allow conjunctive Boolean searches

During the observation sessions it was noticed that there were situations in which preschoolers mimic other children's preferences, but at this stage the egocentrism is prevalent. For this reason it was found more appropriate to have a category named 'favourites' in which children could organize their favourite videos than to have a possibility for them to rate the videos and then present the category 'children's favourites' including ratings from all users of the system, as originally planned. This added feature would also allow some degree of customization, in which children would be able to organize their own category informing the videos they like, giving them a sense of ownership of the system (RL39, REA4).

Children this age do ask for assistance and information to both adults and other children, the way they ask for help or information and how it is provided vary, but the help feature of the system could be inspired on the leading / imitating situations that occur during activities. Preschoolers imitate their friends; a child pretends to be a tiger and roars other children join also pretending to be tigers, roaring and adding tiger movements to the role play. Children also imitate adults, members of staff sing songs and children mimic the lyrics and movements to the song. The help system could then lead children on how to use the application so that they could initially imitate the movements on screen and then use it to scaffold and interact independently (RL9).

Ideally it would be good if the system could offer flexibility for one or more users; in case more than one child are using the system each one could choose a category and then they would watch a video that belongs to all categories chosen. It is, however, a very limited shared experience in which there is only one input device, no room for competing and very restricted way of cooperating. There is no possibility for children to create content and share it with other users either. It was found that to incorporate such features could be interesting, but to provide the possibility for user generated content would be time consuming to both implement and test it. To incorporate in a single evaluation session tasks for children to find video content using the prototype and then produce their own content and share it would be probably not feasible, so these concepts emerged from data did not affect the prototype.

The fact that children sometimes ask for feedback and enjoy reward was taking into account to the design of subsequent sessions involving preschoolers so that they are always complimented and rewarded (RL10).

Two interesting situations during observation sessions had direct impact on the prototype developed. A child was playing with an old keyboard in the garden and kept pressing the keys while saying “broken TV game, broken TV game” (participant 4) and then laughing along with his friends. This data fell under different categories such as multiplayer, inform and role play, but more importantly it indicates that the fact that a keyboard might not be originally and predominantly an input device for the TV, for preschoolers it might as well be. Thus, it was decided to enable interaction with the iTV prototype using the keyboard in addition to the mouse and remote control.

RO4 – Enable interaction via keyboard

The second episode occurred while a participant was watching a friend drawing using a software on the computer. Children had a computer available to play but this was the only occasion during the sessions that participants chose to play with the computer. The child was painting the whole screen so the participant pointed to the eraser and suggested “rub it out” (Participant 4). This 3 year old participant recognised the eraser icon on the screen and was familiar with its function. It was then decided that if there was a need for a button to deselect the filters selected during conjunctive Boolean searches, a ‘clear all’, it would then be represented by an eraser.

4.3.4 Conclusions

The observation offered insight into the age group providing a richer view of preschoolers and their behaviour. Through observation sessions some requirements were reviewed and four additional requirements developed (listed on the Appendix A).

During the process some issues were clarified and fed back to the prototype as design decisions, such as using an eraser as metaphor for a ‘clear all’ functionality.

4.4 Card Sorting Activities

4.4.1 Introduction

The observation sessions did not elicit enough data to establish the categories to be implemented in the prototype. As a result, in order to define and refine these categories it was decided to conduct card sorting activities with children.

Card sorting techniques have been used to create information architecture eliciting conceptual structures from participants in order to reflect how users view the content (Nielsen, 1995). In a conventional card sort activity participants are asked to sort a group of cards with concepts written on them, usually menu entries, categories or headings. After the cards are sorted they are asked to label each group and explain the sorting criteria used. Researchers then study the number of sorts, groups within sorts, the similarity in groups, and the sense of distance of concepts grouped together (Lyman and Lewandowski, 2005).

In open card sorts, recommended for exploration, the number and names of groups are decided by each participant, while in closed card sorts, suggested for assessment, these factors are fixed by the researcher in advance (Hudson, 2005).

Card sorting tasks have been carried out with children as young as eight years of age (Hanna et al., 1999). In this thesis it is discussed the way this technique could be adapted to be used with three and four year olds and contribute to the design of technology for this age group. Three different studies are reported along with advantages and limitations of both open and closed card sort activities.

According to Murray and Reuter (2005), research into children's acquisition of categories helps us understand the impediments to children's use of traditional classification schemes (Murray and Reuter, 2005). Attention to this literature offers an understanding of children's cognitive and developmental needs, which should influence the design of classification schemes and information retrieval systems to better accommodate children's information needs and abilities.

A hundred years of research in developmental psychology suggests that preschool children often categorize by common perceptual properties and/or by thematic relationships (Smith, 2005). Goswami (1998) supports the view that child-basic categories differ from adult-basic categories since children may notice or emphasize different attributes of the same object than adults. Because of their different experiences and their different knowledge, children may then give priority to characteristics rather than defining features in making their category judgments. As a result, categories established by adults may not reflect children's classification; for that reason, in order to create user-appropriate information architecture, is necessary to involve children in the design process.

4.4.2 Closed Card Sorting

4.4.2.1 Introduction

Several categorization activities in psychology studies are based on a task in which children are given a target object and are asked to select the appropriate match from a pair of alternatives. So it was decided to design the first card sorting activity using triads as well, inspired by the Dimensional Change Card Sorting (DCCS) task used to determine extradimensional shifting abilities in preschool children (Kloo et al., 2008).

For the prototype under development, an initial set of categories were pre-established according to a brief analysis of existing applications (see Section 2.3.2). The objective of this study was to check how well the pre-established categories fit children's expectations. It was decided to test the categories: movies, cartoons, music, make and do, animals and nature, fairy tales, superheroes, around the world and TV shows (including live action and puppet shows). It was also decided to test whether children associate video content to the channel which broadcasts it, if so channels would remain as categories. As a result, in addition to sorting a member of each one of the nine categories, children were asked to sort two screenshots members of two different channels, CBeebies and Playhouse Disney, totalling eleven screenshots to be sorted.

In the Dimensional Change Card Sort, inspiration for this activity, children are asked to place cards into trays according to simple rules such as colour and shape. The card sorting experiment described in this thesis, however, requires more abstract categorization ability, involving blurred categories, such as ‘cartoons’ and ‘movies’, that sometimes merge and imposing participants’ understanding that the screenshots presented are representations of video content. Therefore, before confirming if the pre-established categories were appropriate the card sort activity needed to be validated and give evidence that preschoolers can categorize videos based on still screenshots.

4.4.2.2 Method

4.4.2.2.1 Participants

The sample size needed to validate the closed card sorting activity was defined with the assistance of a statistician using the tables for logistic regression and resulted in fifty to sixty children, as already mentioned in Chapter 3.

There were six sessions conducted in five different nurseries. The activity was carried out in a corner of the Nursery setting or in a separate room under the supervision of a member of Nursery staff. Fifty six participants, whose parents have received an information letter and signed a consent form (Appendices D and E), contributed to the study. Twenty seven girls and twenty nine boys, aged between 36 and 56 months, with a mean age 46 months. Most children were from white middle-class backgrounds, but this data was not systematically collected.

Four additional children were excluded from the study; two who inserted all cards in the same box and did not show sufficient understanding of the task, and two due to researcher’s error.

4.4.2.2.2 Procedure

Two side-by-side shoeboxes (Figure 16) were designed, each with a plastic sheet to display a pre-established category plus a slot through which the participant should post a 9cm x 6cm laminated card showing a screenshot from a video. Two predefined categories were displayed at a time and the child's task was to post the screenshot in the box they found most appropriate. For the complete list of the triads, screenshots and the two choices of categories presented, see Appendix F.

There were icons representing each category but the names of the categories were read out for clarity. Some of the screenshots presented were part of a video content while others were from promotion material, such as the one from the movie *Ratatouille* on Figure 16. The videos chosen were children's programmes currently being broadcast and popular children's films and were all strong members of one of the categories presented. The screenshots selected represented the essence of the correspondent video content.

Figure 16. Closed card sorting set up

The triads composed by one screenshot and two categories were kept the same, but participants were randomly assigned to one of three different conditions, each condition with a random order of triads. This randomization was found necessary because on a pilot test most children were mistaken on the three last triads, which could be due to their tiredness at the end of the activity.

The screenshots were printed in colour and the category icons were black and white, to avoid colour associations.

Each session lasted approximately ten minutes and the task was carried out individually by each child.

First the child was asked if s/he watched television and what was his/her favourite programme. Then it was explained his/her help was needed to develop a TV guide for children. The child was asked if s/he could help and it was mentioned that if s/he decided not to help or to stop helping at anytime it was acceptable. The card sorting activity was then described.

The screenshot was shown and it was asked if the child recognised it. If s/he did, the child was asked to post it into one of the two boxes. If the child did not recognise the screenshot, another one of the same category was shown. If the child did not recognise the second screenshot shown, s/he was asked to choose one between the two screenshots of that category and place it in the more appropriate box. Two new categories would then be displayed. In all, eleven screenshots would have been sorted.

The first and main objectives of the task were to check if children were able to relate the screenshot to the video and then categorize the video content. For this reason it would be important that participants were familiarized with the video so they could recognise it from the screenshot and categorize it. Hence it was decided to provide two options instead of one to increase the chances participants were acquainted with the video. The second screenshot was usually part of a video that was also a strong member of the same category that the first screenshot shown belonged but from a different channel, broadcasted at a different time of the day, or mainly directed to a different gender. In case a child did not recognise any of the two screenshots presented it was decided to ask her/him to sort one of them anyway in order to verify if a screenshot could contain enough information to support children's categorization of a video.

After each screenshot had been sorted the researcher said "Thanks" but no positive or negative feedback was given.

When children finished sorting the eleventh screenshot they received a certificate to thank them for their participation.

4.4.2.3 Results

The closed card sorting activity was a binomial experiment in which each trial could result in only one of two outcomes, success in case the child posted the screenshot in the expected box (category which the screenshot was a strong member of, according to adult classification) or failure if the child posted the screenshot in the unexpected box or refused to post it in any box. Success and failure are related to the pre-established categories, in case a significant number of children post a screenshot in the unexpected box this would indicate a failure of the category itself, and that it should probably be modified or eliminated. The term "accuracy" was also used to refer to the degree of correspondence to the pre-established categorization.

Three participants refused to categorize specific screenshots because they were unsure to which category it could belong. Two participants refused to categorize the screenshot from the 'animals and nature' category and one participant refused to categorize five out of the eleven screenshots.

The binomial distribution was used to validate the activity and calculate if the results from the closed card sorting task were above chance. The null hypothesis tested was that the sample data do come from a B (11, 0.5) distribution, implying that the participants are guessing. A probability value (p value) cut-off of 0.05 was then used to reject this hypothesis in favour of an alternative hypothesis that the sample data do not come from a B (11, 0.5) distribution, which imply that participants are able to categorize video content using screenshots and the card sort activity is valid. Since the observed and expected values were very different and the resulted p value less than 0.0005, the null hypothesis was rejected. There is very strong evidence that participants were not guessing and the closed card sorting task can be conducted with preschoolers.

Correlation was used to measure the relationship between the number of screenshots that were posted in the expected category and participants' age. Deak et al. (2002) concluded in their study that children's inference making and similarity selection, such as sorting, are best regarded as "phenomena that emerge under constellations of properties of the child and his/her environment" (Deak et al., 2002). According to the authors preschoolers' matching performance depend on several factors and is difficult to accommodate under traditional theories of cognitive development such as Piaget's that emphasizes age-related changes in static representational structures or capacities.

In this study, from the scatterplot produced (Figure 17), it can be noticed that there is no relationship between the number of screenshots posted in the expected box and child's age. So, within this age group, there isn't any indication that as participants get older they will form categories more similar to adults'. This could mean that other factors also interfere in the categorization ability. Although every participant affirmed s/he watches television, the frequency of media use and variety of content watched, for instance, could have more influence on how preschoolers categorize videos than the child's age.

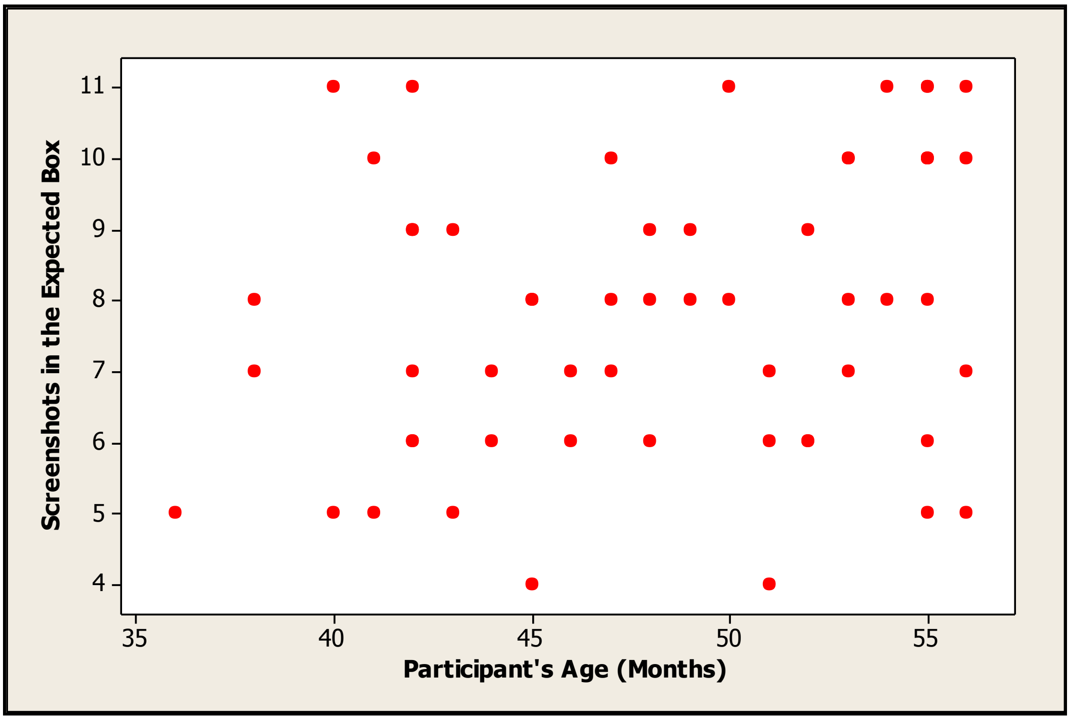


Figure 17. Number of screenshots posted in the expected box by age

The Mann-Whitney test[1] suggests that there is no significant difference between the mean accuracy for boys and girls. Therefore we can conclude that there is not a significant relationship between children's categorization ability and their gender.

A Kruskal-Wallis[2] test indicates that the median accuracy is very similar for the three conditions. Thus, the condition did not significantly affect the results.

---

[1] The Mann-Whitney test is a non-parametric test used to assess whether two independent samples of observations come from the same population. The null hypothesis is that the distributions of the two populations are identical. The alternative hypothesis is that the two populations have dirstributions with different medians (Rees, 2001).
[2] The Kruskal-Wallis test is a nonparametric method to compare medians of populations of more than two independent samples. The null hypothesis is that all samples are from the same distribution. The alternative hypothesis is that at least one pair of medians differ (Sprent, 1990).

Another Kruskal-Wallis test was conducted to verify the impact the Nursery could have on the median accuracy and it indicated that there is a difference in the medians. As it may be seen on the boxplot below (Figure 18) the median accuracy in Nursery 1 appears to be higher than those in Nurseries 2, 3, 4 and 5.

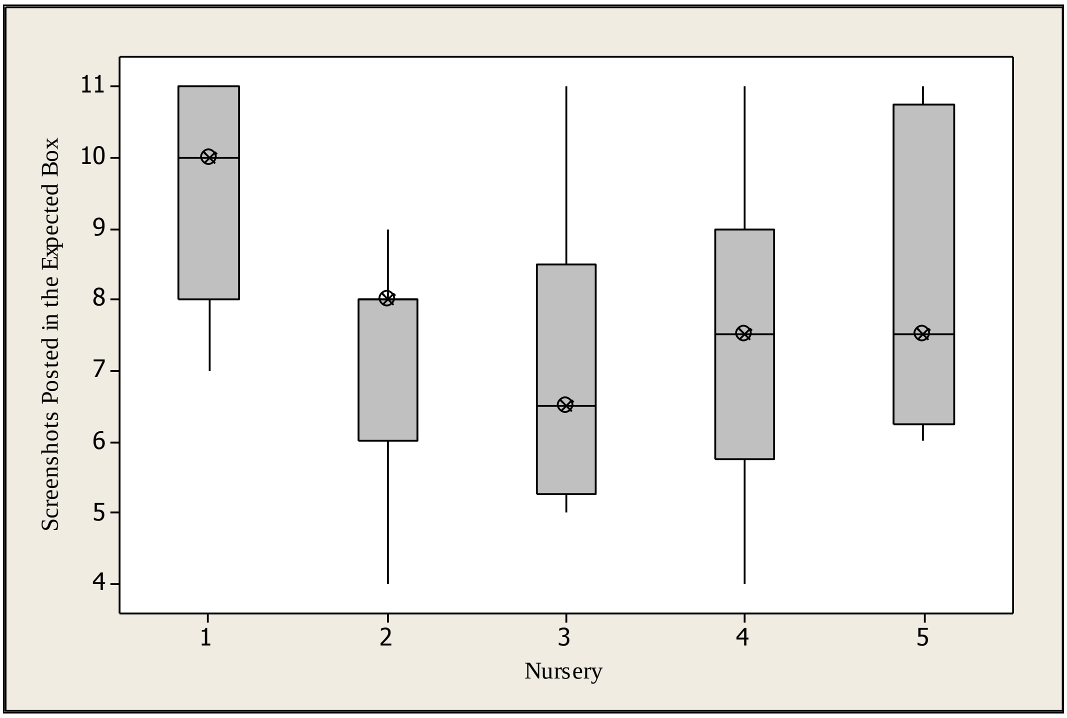


Figure 18. Number of screenshots posted in the expected box for each nursery

This difference could have occurred because the session in Nursery 1 was the only one conducted in a quiet separate room while the sessions in the other nurseries were conducted in a corner of the Nursery setting. The environment noise and activity could have interfered with participants' concentration and influenced the results.

RCS1 - Design and evaluation sessions should be conducted in a quiet room

Most participants instantly recognised the first or second screenshots shown and related it to the video it represented. The only exception was the screenshots from the category ‘animals and nature’. None of the participants recognised any of the videos from either of the two screenshots presented. This was probably because the videos the screenshots came from were not broadcast regularly and not very popular. Nevertheless, despite not being familiar with the video, a significant number of participants posted the ‘animals and nature’ screenshots into the expected box (see Table 1). This indicates that a screenshot can contain enough information for children to be able to categorize its video content.

RCS2 – Screenshots should represent video content on the interface

Given that the results from the closed card sorting task provide evidence that three and four year old children are able to relate a screenshot to the video it represents and categorize it. This activity can be used to assist on the design of technology for young children. The card sorting task was then used to define categories for the prototype application. The results indicate which categories are better understood and which ones may not be as comprehensible for preschoolers.

The numbers were calculated according to the pre-established categories, success as 1, matching the expected categorization, and failure as 0, when children chose the unexpected category or refused to categorize. The mean accuracy for each pre-established category presented in the table below was calculated to support design decisions.

Categories	Mean Accuracy (Standard Deviations)
Movies	0.77 (0.426)
Cartoons	0.68 (0.471)
Music	0.63 (0.489)
Make and Do	0.54 (0.503)
Animals and Nature	0.73 (0.447)
Fairy Tales	0.77 (0.426)
Superheroes	0.89 (0.312)
Around the World	0.59 (0.496)
TV Shows	0.61 (0.493)
CBeebies (channel)	0.80 (0.401)
Disney (channel)	0.84 (0.371)

Table 1. Mean accuracy of screenshots categorized into established categories

The Table 1 above suggests that some categories are understood better than others. Given that 89% of the participants posted the superheroes screenshots into the expected box labelled ‘superheroes’, we can affirm that this category is understood by most preschoolers. Categories like ‘make and do’ and ‘around the world’ seemed to be particularly difficult for children. Considering p=0.5 (50% chance of success and 50% chance of failure), those two categories are just above chance. For this reason they need to be refined or eliminated from the EPG system.

Despite the increasing number of children’s channels available broadcasting the same content and the video on demand features that can disconnect the content from the channel, it was significant the number of children who recognised the channel in which the correspondent video content is broadcasted. As a result children’s channels will be included as categories in the prototype.

RCS3 – Children's channels should be included as EPG categories

4.4.2.4 Conclusions

In addition to the requirements that emerged from the closed card sorting (listed on the Appendix A), some design decisions resulted from this activity. It was decided that 'animals and nature' should be maintained as a category in the EPG system, despite the fact that the screenshots were not recognised, because a significant number of children would identify the category and be aware of the type of its content. Children's channels should also be maintained as categories in the prototype as well as 'movies', 'cartoons'. 'music', 'fairy tales', 'superheroes' and 'TV shows'. Those categories were understood by a significant amount of participants. The 'make and do' and 'around the world' categories, on the other hand, did not appear to be as clear for participants, therefore should be refined or eliminated.

4.4.3 Open Card Sorting

4.4.3.1 Introduction

The categories tested by the closed card sorting were made for children, by adults. There was a concern that child based categories of video content could be completely different from what was pre-established based on existent categorization, so it was decided to experiment other card sorting activities to complement the closed card sorting task.

On the hierarchical taxonomic concept test (Sung et al., 2008) children were presented with four index cards and two baskets and asked to put cards in different baskets according to their categories. After the classification was made a further four index cards would be presented, with two more baskets and children asked to explain the categorization. Initially, the intention was to provide boxes to replicate this study with the screenshots. However, it was decided to ask children to make piles instead and group the screenshots together, otherwise the number of categories that could emerge from the activity would be limited.

4.4.3.2 Method

4.4.3.2.1 Participants

This activity was conducted in one Nursery in Brighton (UK). The inclusion criteria were children's age and a signed consent form.

After the closed card sorting task, children were asked if they wanted to continue to help the researcher by playing another game. Eight participants, seven girls and one boy, aged between three and four years, decided to carry out the open card sorting.

4.4.3.2.2 Procedure

Twenty 9cm x 6cm laminated cards showing a screenshot from a video were used. There were two different conditions. In condition one, children were given five sets of four screenshots each and in condition two children were given four sets of five screenshots each. For a complete list of screenshots given in each set for each condition see Appendix G.

The activity was conducted individually with each participant in a session that lasted for approximately ten minutes.

The activity was briefly explained to the participants as the grouping game, in which they had to put together things that are the same type or kind. The first three participants were assigned to condition one while the five last participants were assigned to condition two. Children were given some time to make groups before the next set was handed. The screenshots in each of the sets were shuffled before being handed to the participant.

Children were motivated to consider all screenshots while making groups, not just the separated sets. And at the end of the activity children were asked to justify their choices.

4.4.3.3 Results

The three first participants put together two different sets of two screenshots, strong members of the same category, from the first set but then struggled to continue the activity when the number of screenshots increased. For this reason the five following participants, were assigned to condition two, in which it was given four instead of five sets and distributed these members of the same category that were being easily sorted to check if this was due to their similarity or because they were both on the first set.

All participants gathered only two screenshots in each group they have made. Most of the screenshots put together were strong members of a category and/or had strong perceptual similarities. The Figure 19 below is an example of cards sorted by a participant.

Figure 19. Cards sorted by a participant during the open card sorting task

Children were asked about their decisions but they could not explain or justify their choices.

There were two screenshots from the same video and although they contained different characters in different scenarios, seven out of eight participants grouped them together. This reinforces the finding that preschoolers are able to relate the screenshot to the video content it represents (RCS2).

Several groups were formed by members of the pre-established categories. Seven children gathered the two screenshots from the category ‘animals and nature’. Six children put the two screenshots representing ‘fairy tales’ in the same group. Three children gathered the two screenshots from videos broadcasted in the same channel. And two children made a group with the two screenshots containing images from music videos, that would be part of the pre-establish category ‘music’.

But non-expected groups were also formed, most of them containing cartoons and animations, including 3D and stop motion. In the pre-established categories of the EPG system the category 'cartoons' excluded animations. Animations were categorized as 'TV shows' if they were TV programmes or 'movies' if they were from films. However, from the open card sorting activity it was noticed that children do not differentiate cartoons, 3D animations and stop motion animations.

4.4.3.4 Conclusions

This activity has to be further developed to help in the design. Children usually become overwhelmed when too many options are shown and cannot associate them, nor can they explain their choices. It may require more of participant's time and would be probably necessary to have several sessions with children to achieve some level of contribution to the information architecture of a system. As a result, the open card sorting could be an interesting technique to be used when young children are technology design partners and contribute to the design process throughout the experience (Druin et al., 1999).

In this research however children are informants (Scaife et al., 1997), they are asked for input at some stages of the design process and play some part in informing the design but their participation is limited as well as the time they collaborate. Hence we found necessary to find another activity to complement the closed card sorting that would not be as time consuming as the open card sorting.

4.4.4 Match-to-Sample

4.4.4.1 Introduction

The results from the closed card sorting indicate which categories are well understood by children and which ones are not as clear. In case most participants relate a screenshot to the expected category the design decision is simple, to maintain that category. However when a category is not comprehended the design decision could be either eliminate or refine the category, and to do so it is essential to identify to which other category children would relate its members. For example, if most children did not insert the make and do screenshot into the expected 'make and do' category box, it demonstrates that the category is not well understood, but it does not indicate to which other category the make and do screenshot should belong.

In order to confirm the closed card sorting results and find if the screenshots that were not inserted into the expected category could fit within another pre-established category instead the match-to-sample activity was developed.

The initial idea for the activity was to ask participants to choose a category between the eleven pre-established ones to paste a screenshot. Nevertheless, this could offer too many options and could overwhelm children like the open card sorting task. So it was decided to base the task on an existent activity to measure children's categorization ability.

Mervis and Pani (1980) used 24 objects forming six artificial categories with four members each. In their experiment five year old children were taught the name of six objects, one in each category. They were then presented with all objects; the researcher would say the name of one object and ask the participant to point or touch all the objects s/he would call by that name. The same procedure was repeated for each of the six category names.

This study was adapted organizing 24 screenshots into six different groups of four screenshots each. The groups were made by some of the pre-established categories and the screenshots to be sorted were some of the ones used in the closed card sorting.

During a pilot study the categories ‘cartoons’, ‘make and do’, ‘superheroes’, ‘TV shows’, ‘movies’ and ‘fairy tales’ were tested. In the pilot test there was only one condition in which children had to paste one screenshot for each one of the six categories. It was found however, probably due to the fact that this was the second activity conducted with each child; participants looked tired after pasting the three first screenshots. So it was decided to ask each child to paste only three screenshots each and assigned participants to one of three different conditions.

Also in the pilot test the open card sorting result was confirmed that most children did not differentiate cartoons from other types of animations. Most children pasted the cartoon screenshot with the TV shows’ screenshots that included animations. As a result, it was decided to test a modified version of the ‘cartoons’ category labelled ‘cartoons and animations’, with one screenshot from a cartoon in condition one and one from a 3D animation in condition three. In addition, the categories ‘movies’, ‘superheroes’, ‘TV shows’, ‘fairy tales’, ‘make and do’, ‘music’ and ‘TV shows’ were tested.

Some of the pre-established categories were left out the match-to-sample. The categories related to children’s channels were eliminated from this task because they seemed to be clear enough. It was the only labelling children did without being requested. During the closed card sorting a significant amount of children related the screenshots to the channel the video is broadcasted, not only while sorting screenshots from channels categories, but when any screenshot was shown several participants instantly identified the channel the videos were from without being asked to do so.

The ‘animals and nature’ category was not included in this activity due to the fact that the screenshots were not recognised by any participant during the closed card sorting, so children would probably not relate the screenshot to the video in the process of categorization as it would be preferable.

The ‘around the world’ category was also excluded because during the closed card sorting just a few children related the screenshots to the videos and the percentage of participants who associated the around the world screenshots to the expected category was low.

On the other hand, most participants recognised the videos from the make and do screenshots. For this reason, despite the fact that a low percentage of participants related the make and do screenshots to the ‘make and do’ category in the closed card sorting, it was decided to test this category in the match-to-sample activity to check if the category should be redefined and/or relabelled or if it should be eliminated.

In the match-to-sample task participants were given a screenshot that was a strong member of one of the pre-established categories, but instead of choosing the category it should belong to, they had to choose the group of other similar screenshots that would be exemplars of its category. The fact that children would not categorize using a category label and icon relating to it, but would be able to choose a group of strong members of the same category would offer a different approach from the closed card sorting that could probably complement the activity.

4.4.4.2 Method

4.4.4.2.1 Participants

This activity was conducted in two nurseries in Brighton (UK), following the closed card sorting session. Like the previous activities, the inclusion criteria were children’s age and a signed consent form.

Following the closed card sorting activity, children were asked if they wanted to do another activity. Fifteen children, eight girls and seven boys, between three and four years old, participated on the match-to-sample task.

4.4.4.2.2 Procedure

Three A4 sheets were used in which were printed in colour 24 screenshots, measuring 3cm x 2cm each, organized into six columns of four screenshots each. Each column was composed by strong members of the same category, based on the pre-established categories. Three stickers, also measuring 3cm x 2cm, in which was printed in black and white, to avoid colour association, the screenshots to be matched, one sticker to be pasted in each A4 sheet.

Children were randomly assigned to one of three different conditions in which they were given three screenshots, one at a time, and were presented with six different categories to match the screenshot. For the complete list of categories and screenshots for each condition see Appendix H.

Each child carried out the activity individually in a session that lasted for approximately five minutes.

Children were given a sticker with a black and white screenshot and asked to put it on the group they thought it would be most appropriate. After the child had chosen a group and had pasted the sticker, s/he was given another A4 sheet with similar categories, but in random order and composed of different screenshots, and another screenshot to be pasted. No feedback was given during the task, but when children were stuck or asked for help the pre-established category of the screenshot to be matched was labelled. If a screenshot from the ‘fairy tales’ category was to be sorted, for instance, in case children needed help the researcher would say the screenshot was from a fairy tale video and would ask the participant to paste it where s/he thought are other fairy tales.

After the three different screenshots were pasted on three A4 sheets children were thanked for their participation.

4.4.4.3 Results

The match-to-sample activity gave some insight on how children would fit screenshots into categories already formed by other screenshots.

As expected from the closed and open card sorting activities children did paste the screenshots from a cartoon and a 3D animation in the ‘cartoons and animations’ category. The sample was too small to generate results that are statically relevant, but four out of nine participants that were assigned to a condition in which they were asked to paste a screenshot from a cartoon or animation included it with the other members from the ‘cartoons and animations’ category, whilst the probability to do so was 1/6, $p=0.17$. This indicates that children may associate both cartoons and animations to the same group.

RCS4 – Cartoons and animations should be included in the same category

The make and do screenshot was not pasted within members of the ‘make and do’ category by any child. Four out of six participants who were asked to paste this screenshot pasted it with members of the ‘TV shows’ category. Children seemed to relate the screenshot containing a person drawing more to other screenshots of live action videos than to screenshots of different arts and crafts from both live action and animated videos. This confirms the findings from the closed card sorting that the category ‘make and do’ is not very clear for children, so the category was eliminated. The Figure 20 below is an example of a ‘make and do’ screenshot in black and white pasted by a participant with members of the ‘TV shows’ category.

Figure 20. Screenshot posted by a participant during the match-to-sample task

Another result from the closed card sorting confirmed with the match-to-sample was the fact that the ‘superheroes’ category is very well understood. The superhero screenshot was pasted with screenshots from the ‘superheroes’ category by five out of six children (p=0.17). And the one child that did not choose the superheroes group pasted it with screenshots from ‘cartoons and animations’, category which the screenshot was also a member.

An interesting fact was that most children that related a low percentage of screenshots to the expected category during the closed card sorting associated most or all screenshots to the ones from the category it is a best exemplar during the match-to-sample task. And the opposite also occurred, children that inserted a high percentage of screenshots in the expected category during the closed card sorting pasted most screenshots in the match-to-sample activity with screenshots from a category it doesn’t belong. This could indicate that children’s categorization abilities could vary significantly, and while some participants could better categorize using category labels others could be best on matching a category member to others from the same group. For this reason it is recommended to conduct more than one type of card sorting activity while working with preschoolers, especially on studies with a small sample.

RCS5 – More than one type of card sorting activity should be conducted with preschoolers to assist structuring the information architecture of a system

As children's abilities to categorize using different methods varies the type of categories they form and use could also differ, not just among the participants but also within children. Nguyen and Murphy (2003) concluded that preschoolers do not rely solely on one form of categorization. The authors examined script, taxonomic and evaluative categories for food items and stated that children can cross-classify items into multiple categories and use these categories for inductive inferences.

4.4.4.4 Conclusions

In addition to the two requirements that emerged from the match-to-sample task (Appendix A), two additional requirements were developed during the activity. It was noted through the task that children are able to cross-classify screenshots based on videos' type (e.g., movies), format (e.g., cartoons) or genre (e.g., fairy tales) suggesting that they are not restricted to a single form of categorization. This indicates that in an EPG application children could benefit from a significant overlap in categories rather than one replacing the other.

Therefore, for the prototype under development it was decided to (1) make each category broader, so it could include all or most participants' grouping choices, and (2) overlap the categories.

RCS6 – Categories should be broad
RCS7 – There should be an overlap of categories

The screenshots pasted within each group of strong members of a category were considered and evaluated how it could become a member of that specific category. This resulted on: ‘cartoons and animations’, ‘TV shows and series’, ‘movies and films’, ‘music and songs’, ‘superheroes and adventures’, ‘fairy tales and fantasy’. In addition, the prototype will also include, as previous established with the closed card sorting task, categories related to television channels aimed at preschoolers.

4.4.5 Conclusions

In conclusion, the closed card sorting could be used with preschool children and give some insight when designing for this age group. The method does have some constraints in comparing just two categories at a time but it can assist in finding when a category is understood.

The open card sorting demands more time. It could be an interesting technique to be used when young children are technology design partners but still needs to be further developed.

The match-to-sample was found really useful especially when combined with the closed card sorting; it can be used when children are informants in the design process contributing to the definition and refinement of categories that would best reflect children’s choices.

The prototype under development incorporated the categories emerged from this process. During the subsequent stages of the research, in evaluation sessions with the prototype, the impact card sorting tasks had on the way children looked for specific videos members of one or several categories is analyzed.

Further research has to be done on the impact these card sorting tasks could have in the design of menu entries and/or headings for different technology aimed at preschoolers.

4.5 Low-Tech Prototyping with Children

4.5.1 Introduction

Previous data collected were informing the prototype being developed but there were still a lot of issues to be clarified such as icons to be used and where to place them on the screen. At this point it was decided to ask children for direct input on the look and feel of the interface. Scaife and Rogers (1999) suggestions for low-tech prototyping with children, such as the use of laminated images which could be manipulated against a background, were combined with some ideas to work with younger children as design partners (Guha et al., 2004) to create a session appropriate for this age group but not as time consuming as the cooperative inquiry.

It is important to note that the card sorting activity was scheduled to take place before the low-tech prototyping sessions, so that the information architecture of the system would be decided and then the icons to represent categories chosen. Nevertheless, following an initial closed card sorting session with a small sample it was decided that the activity could be further explored, and in order to be validated it would need to be carried out with a bigger sample in additional sessions. At that time the low-tech prototyping sessions were already scheduled, so for practical reasons it was decided to conduct the sessions before the card sorting activities. As a result, the low-tech prototyping included categories that would later be eliminated from the prototype. The closed card sorting task, on the other hand, incorporated the icons defined by the low-tech prototyping. And the initial card sorting session conducted before the low-tech prototype sessions was then considered only a pilot study.

4.5.2 Method

4.5.2.1 Participants

Eight children, three and four years old, contributed during this stage of the study. The first session involved four children, two girls and two boys, and lasted 45 minutes, the second session also with four participants, one girl and three boys, lasted 30 minutes.

4.5.2.2 Procedure

During the low-tech prototyping session preschoolers were asked for input and suggestions. Children were told we were working on a “programme finder” and needed their help. They were given a A3 paper in which a TV set was printed, then each one received the first screenshot and asked if they knew the programme presented. After that they were asked to choose one icon among three options provided that would be more appropriate to help children find the particular programme. There were enough options for every child to be able to choose the same icon if needed. Icons not used were collected and another screenshot presented.

The exact same procedure was carried out for nine screenshots, strong members of nine different categories, and then they were asked to choose one icon to assist children find help, one to close or exit the “programme finder”, one to find their favourite programmes and the last one to find games, just to have the option in case it was decided to include games in the prototype. Children were then provided with glue and crayons to create with the material chosen their own “programme finder”. It was decided to include in this activity all categories already considered for the prototype, without taking into account the results from the pilot card sorting activity, so that the findings could be triangulated and reconsidered if necessary.

At the end of the session the prototypes made were photographed so children could keep the prototypes they have created.

The icons and screenshot were coloured, it was considered printing the materials on black and white to avoid colour association. It was found, however, that the colourful images could be more appealing and make the activity more enjoyable for children as well as making the prototype created more similar to a “real” interactive application.

This session involved a large number of materials to be handed in while instructions were given, so it was decided to conduct it with two researchers, one person would explain what children should do while the other researcher[3] would hand the screenshots and icons to be pasted.

The main aim of the session was to obtain information about the icons to be chosen and, if possible, their position on the screen. This could be achieved with the photographed prototypes produced by children, so it was decided not to record the session on video or field notes, this way children would feel more at easy to express their ideas on paper.

The results from the activity are discussed below.

[3] Renata Shimabukuru, a student in the MSc Digital Television Management and Production course at the University of Brighton, conducted her research on iTV for children and assisted in several stages of the work presented in this thesis. She contributed to the low-tech prototyping sessions and helped conducting the card sorting task in some of the Nurseries.

4.5.3 Results

Each child produced a prototype, so there were, at the end of both sessions, eight paper prototypes. The number of children who have chosen each icon provided to relate to the specific screenshot was then calculated, and the icon to represent the category on the prototype decided.

The first screenshot presented was from a children's movie, participants were then provided with three icons to choose to represent the screenshot, a bucket of pop corn, a camera and a roll of film. Two participants chose each one of the icons, so the design decision was not clear in this case, but the three icons could be appropriate.

The second screenshot presented was from a superhero cartoon, three different superhero icons were given, one with an adult like caped superhero flying, another one with an adult like super hero in a leather suit throwing a "fire ball" and finally a child like caped superhero posing. Two participants chose each one of the adult like superheroes and five participants chose the child like superhero, so the icon chosen for the prototype was the child superhero.

The third screenshot presented was part of a music video. Children were given an icon with headphones, one with a character singing and another one with music notes. One child chose the headphones, no one chose the character singing and six participants chose the musical notes. Children were told to choose only one icon, one participant, however, said that on that occasion he would like to use both icons the headphones and the musical notes, because these two icons should be put together. The suggestion was then taken to the prototype.

The fourth screenshot was part of a TV show in which the character and some children were drawing, that would be a strong member of the 'make and do' category. The icons distributed contained a pair of scissors and glue, the other one a paint palette and the third a pot with crayons, brushes and a pair of scissors. Three children chose the pair of scissors and glue, while four participants chose the palette and no one chose the pot with different brushes and crayons. There was however no elucidating detail on the session that would justify maintaining the ‘make and do’ category.

The fifth screenshot was from an animals’ TV show and feature two deer; the icons given to participants were a photograph of a squirrel, a cartoon monkey and a photograph of a dolphin. Two participants insisted on using both squirrel and dolphin, in total three children chose the squirrel, two chose the cartoon monkey and five chose the dolphin. The dolphin was then the icon chosen to represent the ‘animals’ category on the prototype.

The following was a screenshot part of a fairy tale video. The icons participants were given to choose to represent it were a cartoon fairy flying with the magic wand, an illustration of a magic wand, a cartoon fairy posing with pink background. Two girls insisted to have both the fairy and the magic wand, in total two participants chose the flying fairy, five participants chose the magic wand and three participants chose the fairy posing. The magic wand was then chosen as the icon for the prototype application.

The seventh screenshot presented was from a TV show recorded in the Amazon forest, a strong member of the ‘around the world’ category. The icons for this category were a cartoon globe with people from different nations around it, an illustration of the world map and a photograph of a globe. One participant chose the cartoon globe, one the world map and six participants chose the globe photograph. Children, like in the card sorting activity, during the low tech prototype sessions did not recognised the ‘around the world’ screenshot, and there was no information gathered in this session that would justify maintaining the category.

The eighth screenshot was from a popular children's TV show, children instantly recognised it and were offered the chance to choose a cartoon TV as an icon to represent it, a photograph of a TV (with a blue screen) and a second TV photo with some curtains and holophotes on the screen. Two participants chose the cartoonish TV, four chose the TV with no content on screen and one chose the second TV with an image of red opening curtains on screen. The TV set with a blue screen was then chosen as the icon to represent the 'TV shows' category of the prototype.

The last screenshot was from a cartoon, also popular and immediately recognised. Children were given a blue cartoon "splash" icon with a smiley face, a purple cartoon monster holding a flower and a cartoon little character on a background containing a rainbow, sun, clouds and mountains. Six participants chose the last colourful icon, one of the participants mentioned he would not choose because none of the icons supplied could represent the cartoon screenshot presented. This should not be ignored because it was the only situation during the activity that a participant justified not wanting to choose an icon. It is certainly difficult to represent on a single icon the entire 'cartoons' category, so the little monster on the rainbow background was chosen but could be provisional in case a more appropriate icon was found in later stages of the design.

Children were then asked if users wanted to access help which icon would be appropriate, no one chose a question mark, three participants chose a lifesaver and four participants chose an illustration of an adult holding a child's hand, that was then the option considered for the prototype.

Participants were told to choose an icon to exit the application or close it, no one chose the illustration of a door, two participants chose the illustration of a man exiting through a door (part of the emergency exit signs in the UK) and four participants chose an icon of a white "X" on a red background (usually used to close computer programmes), chosen then to be included in the prototype.

After that, children were asked to choose an icon to find their favourite programmes, no participant chose the yellow star, four participants chose a red heart and three participants chose and exclamation mark. The heart was then chosen for children to classify and find their favourite programmes on the prototype.

The last request was for children to choose an icon to find games; it was not yet clear at this stage if games would be included on the prototype so it was found that it could be useful to include it in the session. Three options were provided, two children chose an illustration of two dices, no one chose an illustration of a ball and a toy car, and four children chose a photograph of a video game joystick.

The positions icons were placed on the paper prototypes varied significantly and could not be interpreted to inform the design of the prototype. Figures 21 and 22 are examples of low-tech prototypes created by two participants.

Figure 21. Participant concluding her low-tech prototype

Figure 22. Participant explaining his low-tech prototype

An unexpected result from the session was the way children went beyond the choice of icons and their place on the screen and explained how the interaction would take place. A boy during the first session (participant 3) elaborated on how people would use the buttons (referring to icons provided) and what would happen on screen when the buttons were pressed, to express his ideas on the paper prototype he glued and removed icons and screenshots while talking about the actions. A girl (participant 6) related the action on the remote with the action on the screen and said there could be an “X” button on the remote to activate the “X” button on the screen. Unfortunately, because the session was not recorded, details such as words children used to express themselves were lost.

4.5.4 Conclusions

As the findings from the card sorting activities indicated, the images from both marketing materials and screenshot are easily recognizable and related to the video they represent (RCS2). Thus, such images may be used in the prototype application interface to assist children on finding the respective videos. Movies, for instance, could use the marketing images, while TV shows and cartoons could make use of screenshots representing the essence of each episode.

Results did not offer any insight on the position on the screen the icons should be placed, but the sessions were incredibly useful to determine the icons to be used on the interface. It was not so straight forward as planned in a way that children sometimes asked to have more than one icon to represent a category. There was the initial rule that they should choose only one icon, but flexibility was important to maintain the session fun, and in the occasions participants did justify the reason they wanted more than one or none of the icons it was very fruitful.

The fact that some children talked about the interaction process while dealing with the paper prototypes, indicates that more interactive materials, such as Velcro instead of glue, should be provided on low-tech prototype sessions, so that they can have structure to elaborate on ideas. It is also recommended to video record the sessions to analyse in detail children's explanations about functions of their prototypes.

RLTP1 – During low-tech prototyping sessions interactive materials should be used

RLTP2 – Low-tech prototype sessions should be videotaped if possible

These are then two requirements for future low-tech prototype sessions that are not part of the study reported in this thesis, but could be useful for future work.

4.6 High-Fidelity Prototype

High-fidelity prototypes are prototypes produced with the final product materials and look like the final thing (Preece et al., 2002).

At this stage on the prototype development it was decided to produce a high-fidelity prototype to clearly define the navigational scheme, the look and feel of the final prototype and to test such features before the evaluation sessions.

This version of the prototype was inspired on children's toys like mini-laptops (such as the one presented in Figure 23) and portable video games like Nintendo DS (Figure 24) that have games released specifically for the preschool audience.

Figure 23. VTech's Mylaptop ©　　Figure 24. Nintendo DS ©

The inspiration taken from these examples was the concept that the content presented on the screen is altered by pressing physical buttons. It was decided to make the interface so that buttons, imitating physical buttons (RL24), are used for interaction with the prototype and when pressed alter the content presented on the screen. It was also decided that the interface would open and close like these devices providing additional association with physical toys children are familiar with.

It was important to base the prototype concept on such technology familiar to children to aid them to develop an appropriate mental model to operate the system. Norman (1983) defined mental models as evolving models that people formulate through interaction with a system. The purpose of seeking inspiration for the prototype in technology preschoolers have access to was that children could apply the abstractions developed on how these systems work to the prototype. They will then be able to use their previous experience with similar systems to develop a mental model that matches the conceptual model of the prototype application. As a result, they may interact with the system intuitively.

Inspired by the examples above (Figure 23 and 24), based on requirements gathered thus far, the prototype wireframe (Section 4.2) and on the results from all design activities carried out to this point, the high-fidelity prototype was developed (Figures 25 and 26). For the screen interface to mimic a physical device it was modelled and animated using the software 3D Studio Max 8. The prototype programming was done with Macromedia Flash 9.0.

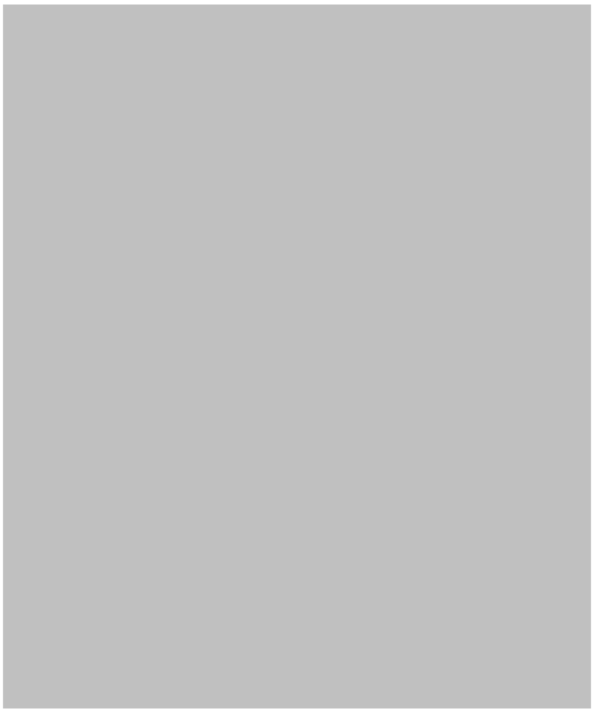


Figure 25. High-fidelity prototype main menu

The prototype menu is composed by purple buttons containing categories, screenshots and coloured buttons with functionalities such as exit and help. All items in the interface have at least 64 pixels diameter (RL21) and were placed close to each other but distanced enough to compensate inaccuracy in targeting (RL18).

Only children's videos are available to be accessed through the interface (REA5). Children may search and browse video content (REA1) via pre-established categories (REA2). The categories were decided with children during the card sorting activities to reflect their concept of categories (RL5), they are: 'cartoons and animations' (RCS4), 'TV shows and series', 'movies and films', 'music and songs', 'superheroes and adventures', 'fairy tales and fantasy'. In addition to those there were the categories related to television channels aimed at preschoolers (RCS3): 'CBeebies', 'Playhouse Disney', 'Nick Jr.' and 'Cartoonito'. No text is used in the interface (RL6); the categories and functionalities are represented by icons chosen during the low-tech prototype session.

The category represented by the heart icon is the 'favourites' category composed by videos chosen by the user. Children can rate a video as favourite when watching it and the video will then be easily accessed through the 'favourites' button on the menu. It was found to be a way to support customization of the system (RL39) and to provide opportunity for repetition of content (RL38).

The navigation works as previously described with the prototype wireframe. To choose a category, users may click using the mouse or press the OK button using the remote control on the purple buttons containing the icons that represent each category. The screenshots from the chosen category are then displayed in the 3D wheel on the upper screen. To assist users on the navigation there is a visual feedback highlighting the category selected and an audio feedback consisting of the category label (RL23, RL35 and RL41). The cursor activation area is slightly enlarged though the visual feedback to enhance target acquisition (RL22).

The system adopted a faceted approach (REA3) in which users may create conjunctive Boolean searches (RO3) by selecting more than one category.

The screenshots are the most important items of the interface, because they represent videos (RCS2) and give access to the video content, therefore they are presented in the middle of the screen (RL26). The screenshots move around the wheel automatically. Users may speed the wheel movement by bringing the mouse to one of the wheel's corners or by using the arrow keys on the remote control. It was decided to eliminate the arrow icons previously presented next to the wheel to reduce the number of items on the screen, but to leave the function to be used by advanced users providing flexibility and efficiency (RL29).

Colour buttons on the remote control may be used as short cuts (RL2). From the main menu, when users click on the green 'X' button or press the green button on the remote they exit from the application, and the prototype closes (Figure 26).

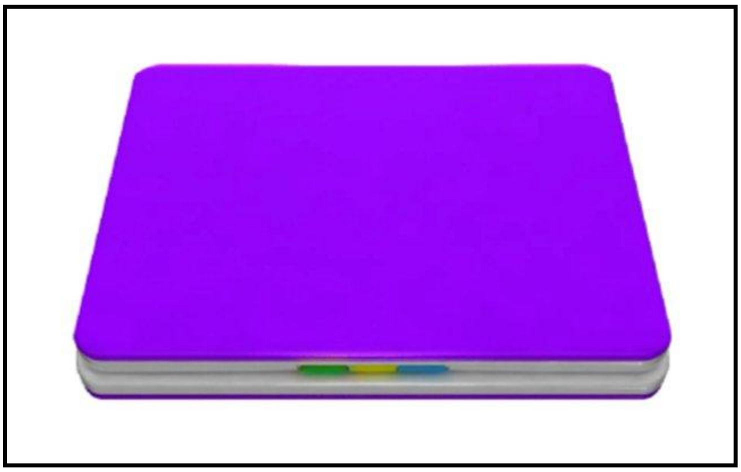


Figure 26. High-fidelity prototype closed

In order to deselect all selected categories children may click on the yellow button with an eraser icon or press the yellow key on the remote. To access the help section with information about how the system works users may click on the blue help icon or press the blue button on the remote control.

To access the video segments users may click with the mouse or press the OK button with the remote on the screenshot in the 3D wheel representing the video they would like to watch. The video selected is then presented in full screen (Figure 27).

Figure 27. High-fidelity prototype video playing

It was found useful to have the three buttons with functions such as exit and help on the top of the interface so that when the video was played, and the category buttons not needed, the screen could be zoomed in (Figure 27) and these buttons would still be presented consistently (RL27) in the same position. At this stage, however, the ‘clear all’ icon had no function and was replaced by the icon to rate the video as ‘favourite’ represented by the heart icon. In addition, the exit button in this screen directs the user back to the main menu, instead of closing the entire application, as in the main menu.

It was found problematical to insert humour in an EPG interface because it could become boring for an application used frequently and could interfere with the navigation. For this reason, it was decided to add a subtle funny noise to the animation of the application opening and closing. This way there is a somehow entertaining characteristic to the system (RL11) that does not annoy the user or gets in the way of the EPG main functionality (RL28).

The prototype colour was chosen to be purple because during previous sessions with children it was found to be a colour that pleased both boys and girls. The prototype was at this stage named ‘Purple’ after its colour.

Following the development of this high-fidelity prototype it was decided to evaluate it focusing on the redesign, so that the evaluation sessions could be carried out with a final prototype containing fewer problems.

4.7 Expert Evaluations

4.7.1 Introduction

Expert evaluation is an inspection method that relies entirely on experts’ judgement, with no user involvement. There are several types of expert evaluation methods such as reviews, heuristics evaluations and walkthroughs that lead experts on the inspection of an interface to predict problems users would have to interact with it (Preece et al., 2002). Expert evaluations can be used in any stage of a project, in some cases they are an alternative for user testing sessions, but on the work reported in this thesis, experts from academic and industry were asked to analyse the first working version of the prototype and make suggestions for improvement so that the prototype could be improved and shown to children on subsequent evaluation sessions with a reduced number of navigational problems.

4.7.2 Method

4.7.2.1 Participants

Nielsen and Molich (1990) recommend that heuristic evaluations should be done with three to five experts, they stated that there is no gain on additional information by using larger number of evaluators. Research involving expert evaluations for technologies for children, however, have indicated that the number of evaluators should be significantly higher (Bekker et al., 2008).

In sum seven female, seven male participants, hereafter called experts, from twelve institutions in seven different countries evaluated the prototype. They were experts in one or more of the following fields: children, technology for children, human-computer interaction and digital television.

A study group on human-computer interaction from the Federal University of São Carlos (Brazil) also evaluated the prototype, they analysed it in group but submitted only one evaluation form, therefore they will be considered during the analysis as one additional expert.

Evaluation forms were sent via e-mail to the nurseries that had contributed to the previous stages of the study; they were, however, not returned. For this reason, there were no experts specialized on children only, but there were five experts in technology for children and three out of these five professionals were specialists on children and technology for children, including experience on child psychology and teaching. In addition, among the experts, there were thirteen professionals with expertise on human-computer interaction and seven with experience on digital television.

Their level of expertise varied from one to five years to more than fifteen years. Eight experts had one to five years experience in the field, two experts had five to ten years of experience, two experts had ten to fifteen years experience and three experts had more than fifteen years of expertise in their field.

4.7.2.2 Procedure

Evaluation forms were e-mailed for experts and others were completed during the Euro iTV conference held in Salzburg in July 2008.

In the form the prototype was explained and instructions were given to complete the evaluation, experts were asked for their field of expertise and years of experience and then a set up information was included for those experts who had the form e-mailed to them.

First experts were asked to conduct a cognitive walkthrough (Wharton et al., 1994), in which they should have performed a task and checked for each step how easy it would be for a new user to accomplish the task. Although some studies defend that the cognitive walkthrough can be a reliable source to indicate usability problems on interfaces for children (Campos and Mano, 2006) other researchers indicate how complex the process can be, since in order to conduct a cognitive walkthrough for a children's application the expert would need to be able to think like a child (Read, 2005).

So it was decided to combine the cognitive walkthrough with the structured expert evaluation method. The walkthrough would be used as a way to provide structure for experts to explore the prototype, and then they would be asked to answer some questions with their opinion about the overall system. These questions were based in the structured expert evaluation method (Baauw et al., 2005), an analytical evaluation method designed to assess fun and usability of young children's computer games, and adapted to suit an iTV application. The complete form sent to the experts can be found on the Appendix I.

The evaluations assisted in improvements on the prototype and together with the literature review also helped to stimulate thinking about properties and furnish initial ideas to be used in the design principles.

4.7.3 Results

Most experts highlighted the comments and suggestions given during the cognitive walkthrough on the following Structured Expert Evaluation (SEEM) questionnaire, but the results from the two methods will be firstly presented separately to clarify the way one method complemented the other.

During the cognitive walkthrough, experts' concerns and suggestions were divided into three main categories: icons' representation, feedback and help features.

There were several issues raised according to the icons' representation on the prototype. Seven out of the fifteen experts mentioned the icon representing 'cartoons and animations' was difficult to be identified and two experts stated that the 'music and songs' icon was not clear. One expert pointed out preschoolers would not comprehend children's channel textual logos used as icon. Ten experts were concern with the exit icon. Six stated the icon could be hard for children to recognise. Six experts suggested having two different icons one to go back and another one to exit the application, instead of one icon being used for both actions in two different instances. And four experts recommended the icon should be coloured red instead of green. Two requirements emerged from these suggestions.

RE1 – Two different icons should be presented, one to go back the other to exit the application
RE2 – The exit icon should be red

Help features were also pointed out on the cognitive walkthroughs forms, one expert said the help narrative was too long and suggested dividing it on contextual help sections and another expert recommended a tutorial segment to run as soon as the application starts including the information accessible via help.

RE3 – Instructions and help section should be divided into small segments
RE4 – Tutorial should be provided for inexperienced users

According to the feedback provided during interaction using this first version of the prototype application, six experts stated that the visual feedback was too subtle therefore it was hard to differentiate when a button was selected, on or off. Most experts recommended the feedback should be emphasized, three experts suggested displaying the icons from selected categories on the screen and one expert suggested having different audio feedbacks in distinctive instances of the button. Two experts mentioned the rollover audio on buttons when selected sometimes overlapped and both suggested inserting a 0.5 of a second delay before the audio feedback is played.

RE5 – Visual feedback should be prominent
RE6 – 0.5 second delay should be added to audio feedback

In addition, one expert raised the concern that it was not clear the screenshots on the prototype screen were selectable and recommended sound to be added to the screenshots as well. Two experts criticized the video loop; one of them suggested inserting buttons to pause the video being watched.

RE7 – Video loops should be removed

RE8 – Users should be allowed to control the video

Experts' responses using SEEM indicate that children will understand the system and will be able to interact with it. According to most experts, preschoolers will be able to perform physical actions needed and will enjoy the search process.

The concerns and suggestions rose using SEEM were similar to the ones presented on the cognitive walkthrough. Some experts highlighted their recommendations, while others explained issues already mentioned by other experts during the walkthrough. It can be noticed; according to the graph below (Figure 28), that the main problems identified through SEEM not previously found with the cognitive walkthrough regarded the representation of icons for categories and functionalities such as 'clear all' and 'help'.

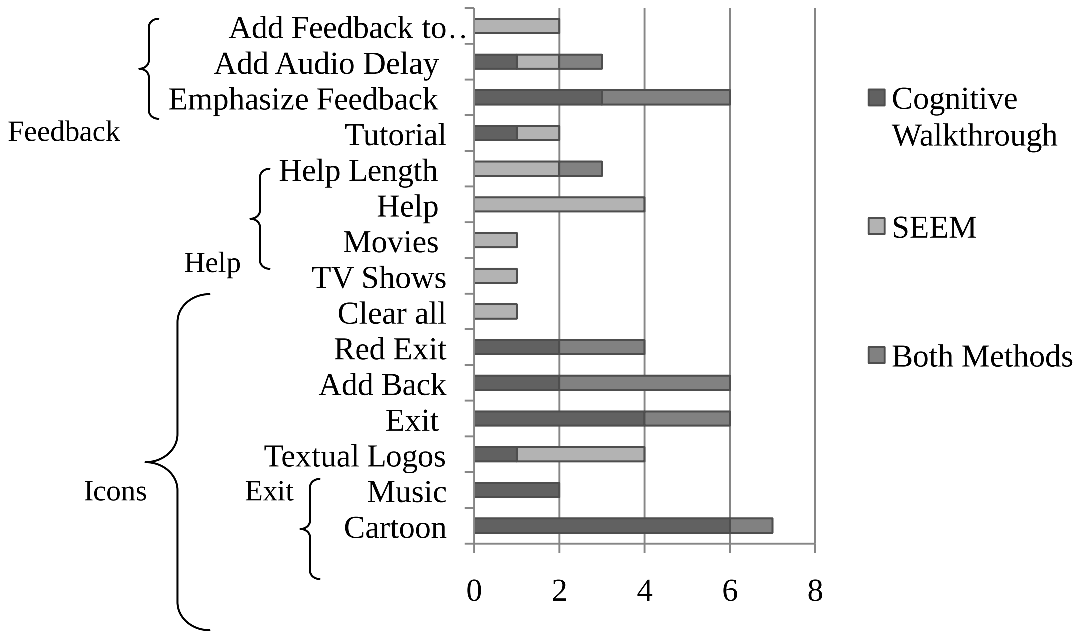


Figure 28. Number of problems identified by experts using each method

Other comments from experts included the fact that the interaction via keyboard should be enabled, to assist children who may have difficulty interacting using the mouse (RO4). In addition, two experts found that the three coloured buttons on the top of the prototype were too small and close together for children to click.

One expert pointed out that the name of the prototype could confound children during interaction, because they have to press the green button, for instance, to go back to 'Purple'. For this reason the prototype should not be named after a colour.

4.7.4 Conclusions

The cognitive walkthrough and structured expert evaluation methods certainly complemented each other, contributing to a holistic analyse of the prototype and to the development of eight requirements (Appendix A).

All experts' suggestions were considered, some had to be discarded for various reasons. For example, in order to add audio feedback when the button is inactive justifying the inactivity, e.g. there are no superheroes cartoons on the CBeebies channel, would mean providing a long audio feedback that could confuse the user instead of explaining the situation. According to requirements from literature icons should be associated with simple words (RL23). To enable audio feedback on screenshots representing videos was found problematic as well, in case different episodes of the same programme were available the audio feedback would then have to include the episode name and this would also result on a long audio feedback. The fact that in previous stages of the research children had no difficulty identifying the screenshots and relating them to video content indicates there is no need for audio feedback on screenshots.

Some experts suggested including the icons from the categories selected on the screen emphasizing the feedback. The International Children's Digital Library implemented something similar, an equation that presents the user with the categories selected during the search. Preschoolers, however, would not understand the equation, and it was found that adding an icon on screen similar to the one on the button they just pressed could lead the child to assume it was another button, for this reason this suggestion was discarded.

On the other hand, several experts' suggestions resulted on modifications implemented on the prototype.

The highlighting was emphasized to give children more feedback during interaction. The 0.5 second audio delay was inserted before the feedback (RE6). The arrow keys on the keyboard were implemented as an option for children who found it difficult to use the mouse. This requirement was already indicated during the observation session (RO4) but not yet implemented on the high-fidelity version of the prototype.

It was decided to start the evaluation sessions with the help as a tutorial for children to interact (RE4). The help section was divided into two contextual fragments (RE3), when accessed through the main menu the child would be able to view information about the interaction with the menu and when accessed during the video the child would be able to see how to interact on that screen, save the video as a favourite and exit.

The loop was removed from videos (RE7). As soon as videos were finished the prototype would take the user back to the initial menu. A pause/play button was also added during the video for children to have control of what they were watching (RE8).

There were, however, issues to be considered that experts rose as problems but no viable solution was proposed. The crucial problem that the larger number of experts identified was the cartoon icon. Some ideas were proposed to add a cartoon character, for instance, but this would imply in choosing an unknown character that children could then relate to most cartoons. A very good recommendation to solve this problem was to add photos for all the other icons instead of illustrations, this way it would be clearer that the only icon that was an illustration was related to the category ‘cartoons and animations’. This was, however, discarded because the icons were too small to enable differentiation between photo and illustration. The ‘TV shows’ icon, for instance, was already a photo that the same expert recognised as an illustration. So it was decided to ask children for assistance to choose an appropriate cartoon icon.

It was also decided to ask for children for input on the icons (and their colours) to be used to go back to the menu and exit the prototype, to be then also decided if it should be the same icon or two different ones.

In the same session with children, it would be verified if it was clear that the screenshots were selectable and if the coloured buttons were actually small or too close for children to click. These issues could have been modified according to experts’ suggestions, by making the screenshot look selectable and enlarging the coloured buttons, but at the same time it was found important to implement as many experts’ ideas as possible to improve the prototype it was also decided to ponder the decisions and leave some of the recommendations, especially those suggested by a small number of experts, to be tested by children. Following the existing guidelines, all buttons on the interface had at least 64 pixels diameter that according to Hourcade et al. (2004) is enough for children as young as four years old to click (RL21), so the buttons could be small for younger children but only through user testing it could be identified if such problems affect users’ interactions or if they are only false alarms (Hartson et al., 2001).

4.8 Prototype Adjustments with Children

4.8.1 Introduction

As mentioned before, there were some problems raised by experts with no clear solutions, concerns, for instance, that some of the icons could not be easily recognisable and other issues, such as the size of the colour buttons that had to be verified. For this reason it was decided to involve children in a session to adjust the prototype.

The icons used on the prototype were chosen with children during previous stages of the research, but they were provided with a very limited amount of options. So it was decided to test if the icons that were most criticized by the experts could be replaced by a more meaningful option.

The main goals of this session were then to redefine some of the icons and verify children’s abilities to interact, so it was decided that a high tech prototyping session or agile development, in which children's suggestions would be fed back to the prototype that would be re-coded to be re-tested later in the same day, would not be necessary. Only some adjustments were needed so they could be made instantly during the session.

4.8.2 Method

4.8.2.1 Participants

Four children, one girl and three boys, between three and four years old, participated in this session. An additional child also took part in the study but the data was lost due to technical problems. These children had already been involved on one or more previous sessions.

4.8.2.2 Procedure

In this session, children were asked to assist on the development of the TV finder, they were explained their participation was voluntary and they were allowed to leave the session in case they wanted. After that children were shown the application help, the tutorial recommended by experts (RE4), which explained how the prototype worked. Then they were asked to perform the same task as the experts, however they were not told the exact steps to be followed, so tips were provided if they got stuck.

As soon as they accomplished the task another version of the prototype in which other options for icons were included was shown. Children were asked if they could identify the icon for a certain function, then they were shown two other options and asked which one amongst the existent and the two other options was the most appropriate icon. The optional icons could be dragged and dropped above the existent one replacing it.

The activity was conducted individually with each participant. The data was video recorded for analysis.

The session had a clear structure and was conducted with a very small sample, so it was found there was no need for a memorized script this time.

This activity could probably be related to a high-tech prototyping session, yet naming it prototype adjustment with children was found more appropriate because, as already mentioned, it was decided that at this stage the only crucial input needed was to check the appropriateness of certain icons and verify if children understand the screenshots are selectable and could use the coloured buttons, so participants were not given enough room and structure to opinion.

4.8.3 Results

During the session it was noticeable that there was in fact a severe problem with the cartoon icon, when asked among the categories' icons on screen which one was representing cartoons, no child recognised it. Nevertheless, there was no clear substitute for it. Two children chose an image with several cartoon animals, one child chose a cartoon dog and one child chose the original icon.

On the other hand, the exit button (green 'X') was very clear. After watching the video segment participants were asked if they knew how to go back to the menu, two participants instantly clicked on the exit button, while two participants mentioned they did not know and were shown the button on the screen. Later on the session, when children were requested to choose between icons, they were first asked to identify the icons on the screen; at this stage all four participants indicated that the green 'X' was the exit button. Subsequently, three out of the four participants agreed that the icon should be the red 'X' (RE2), one participant chose the green 'X' originally presented on the interface, and no one chose the image of the prototype closed. It was then decided to use the red 'X' icon on the prototype.

According to one participant the back and exit icons should be the same image represented by the red 'X', the three other participants chose different icons to represent back and exit (RE1). Two children chose an image of the prototype menu to go back and one participant chose the back arrow.

Regarding the interaction process, there was no indication whatsoever that the coloured buttons were too small or close together, participants had no difficulty on clicking on these buttons. Children had no problem identifying that screenshots were selectable either, the four participants clicked on the screenshot without being told to do so, it was probably intuitive because one participant clicked on the screenshot from his favourite video as soon as the prototype application was opened even before watching the tutorial.

During this session, like on the closed card sorting activity, it was also noted that children recognise the channels' textual logos and relate the videos to the channels they are broadcast. A channel button with a textual logo was the first selected by two participants when asked to find the video.

One of the participants chose to use the computer touch pad instead of the mouse; this indicates the need for flexibility on input devices to cater for users with different abilities and preferences.

RPA1 – Enable interaction via touch pad

Participants did not have much room for exploration during the session and there was only one video segment available to watch, this generated some frustration. Some children clicked on several screenshots and mentioned the videos they wanted to watch and it had to be explained that at this point there was only one option available.

4.8.4 Conclusions

Only one requirement emerged from the prototype adjustment sessions: enable interaction via touch pad. Nevertheless, several design decisions were resulted from this activity.

Following the analysis of the results from the prototype adjustments sessions, the icon with cartoon animals was chosen to represent ‘cartoons and animations’. It was however found, after rendering the image, that the animals were too small to be recognizable on the label button, so the original icon was maintained.

It was decided to differentiate exit from back (RE1). The red ‘X’ chosen by most participants was then selected to exit the application and the arrow chosen to go back to the menu. The decision for the red ‘X’ was clear, most experts recommended and three out of the four participants agreed it was the best option. In fact, it is worth reporting, that during the low tech prototyping session the cross presented as an option to exit was actually red, but because in the United Kingdom the red button is used to enter interactive services it was decided to use a green version on the high-fidelity prototype. After experts’ reviews and this session with children, however, it is noticeable that the fact that the red button is used to access interactive services on the TV do not overcome the widespread association the colour red has with exit or close. Before the button was named exit, mainly because it was related to the emergency exit colour in the UK, with the alteration for red it was renamed to close, related to the feature of most computer programmes and to the interface action, as the close button is pressed the prototype would then close.

The arrow was chosen as the icon to lead users back to the menu, despite the fact that the image of the prototype had more votes, in order to maintain the consistency with the computer interface, along with the red ‘X’. In case the image of the prototype closed was chosen as the option to exit/close the application then the image with the menu would be used to go back.

The coloured buttons and screenshots were kept the same as there was no indication they had to be altered to improve interaction.

The fact that this session was conducted with a very small sample indicated that the results had to be considered, but the modifications to be implemented needed to be determined carefully. Some design decisions were very clear, such as differentiating exit from back (RE1) and making the exit icon red (RE2). Other decisions, however, such as the icon to represent back and the icon for the category ‘cartoons and animations’ could not be only based on children’s choice, since the numbers were too small and the icon with more votes had the advantage of only one child.

4.9 Conclusions

The activities reported in this chapter were very useful as they originated a list of requirements presented in the Appendix A.

The requirements are the main contribution from this chapter as they provided a basis for the design of the prototype application and most importantly, after validation during evaluation sessions, will contribute to the framework and be translated into design principles.

The next chapter of the thesis reports the evaluation sessions conducted with the prototype developed through the activities presented in this chapter. In Chapter 6 the framework is presented together with design principles and a review of refined methods for design and evaluation with young children. In Chapter 7 conclusions are presented with ideas for future work.

Chapter 5. Evaluation Methods and Results

5.1 Introduction

In the previous chapter, the design activities that contributed to the development of the prototype application were explained, requirements established and design decisions made. This chapter describes the process of evaluation of the prototype. Evaluation sessions are described in Section 5.2; results are reported in Section 5.3; and the effect that requirements and design decisions had on participants' experience is discussed in Section 5.4. In the next chapter these findings will contribute to an understanding of children's interactions with iTV, assist in constituting the framework and supporting design principles.

5.2 Method

5.2.1 Participants

It was decided that a small sample would be appropriate for the evaluation process, as explained in Chapter 3 (Section 3.4.3), so that children's interactions could be analysed in depth to yield insight into of a phenomena not yet well understood.

The sample was selected from two different countries, the United Kingdom and Brazil, to provide diverse conditions and variations with the aim of enhancing the explanatory power of the study.

There were four sets of evaluation sessions conducted in four different nurseries: two nurseries in São Paulo, one private and one public, and two nurseries in London, one voluntary and one private.

The activity was carried out in a separate room under the supervision of a member of Nursery staff (RCS1). Twenty two participants contributed to the study, eleven participants in Brazil and eleven in the UK, twelve girls and ten boys, aged between 37 and 59 months, with a mean age of 48.7 months. Children were racially and culturally diverse and from various family backgrounds, but this data was not systematically collected.

Eight additional children were excluded from the study. Three participants were outside the age group being studied. One participant in Brazil was six years old and two participants in the UK were two years old. One participant was excluded due to researcher's error. In this session the computer screen was positioned incorrectly so that the webcam only captured the participant's forehead. Four participants were excluded due to technical problems in which one of the software tools being used to record the session failed. In two sessions the webcam did not capture the entire session. In one session the screen capture software failed and in one session the software recording the clicks and buttons pressed failed. A number of these technical problems were caused by software malfunction while others were caused by participant's interactions; some children held the remote control over the keyboard, occasionally resting it on the keyboard and accidently pressing keys that were equivalent to short cuts which closed the programmes running behind the prototype.

5.2.2 Materials

5.2.2.1 Consent Forms

The University of Brighton Ethics Committee and the Nursery managers approved the study. The parents of participating children received a letter explaining the research (Appendix J) and signed a consent form (Appendix K) authorizing their children to take part. Children were asked for consent verbally. Before the start of the session the study was explained to the children, they were then asked if they would like to participate. The evaluation session was carried out only with children whose parents had signed the consent form and who agreed in taking part themselves.

5.2.2.2 Parents' Questionnaires

A questionnaire was used to gather information about participants' television viewing habits, computer usage and previous experience with remote control and mouse. Parents were asked for an estimate of the frequency of their child's TV viewing. The categories recorded were; several times per day, about once a day, a couple of times a week, about once a week, about a couple of times a month, about once a month, less often, never or didn't know. Then they were asked if their child used the remote control while watching television, did not use the remote control or only used with supervision. Both questions were also asked regarding participants' computer and mouse usage. A copy of the questionnaire is attached as Appendix L.

As questions were based on questionnaires used in published studies involving hundreds of participants, it was considered unnecessary to pilot test the questionnaire. The questions about the frequency in media usage were inspired by the Livingstone's (2004) research UK Children go Online, and the question about input device usage was based on Hynd's thesis (2006).

The main aim of the questionnaire was to identify participants' media and input device usage in order to relate it to the experience they had while interacting with the prototype application.

5.2.2.3 The Prototype Application Tested

The evaluation sessions were conducted with the prototype developed during the previous stages of the research, discussed in Chapter 4. Before the prototype was tested, however, some modifications were made considering the requirements that emerged from the expert evaluations and prototype adjustment sessions with children. In addition, the prototype was translated to Portuguese to be evaluated in Brazil.

The interface of the prototype application tested, as the high-fidelity prototype, was composed of buttons representing functions and categories, and screenshots representing videos.

According to experts' suggestions and results from the prototype adjustment sessions two different icons were presented; one to go back, the other to close the application (RE1). The 'back' icon was illustrated by an arrow and the 'close' icon by an 'X', both white in a red background (RE2). The audio feedback, previously 'exit' in both instances, was changed to 'back' and 'close', but both the 'back' and 'close' buttons would still serve as escape route (RL36).

Considering the feedback, a 0.5 second delay was added to all audio feedbacks in the interface (RE6). The visual feedback was enhanced so that there is a light animation over the button on roll over, when the categories are selected (RL19), and when the categories are chosen the buttons light up (RE5).

Following a suggestion from an expert the name of the prototype was changed from 'Purple' to 'Zap'. In theory an EPG could be used to replace zapping, that means switching channels in both Portuguese and English, so the prototype could be named after this feature. The term may not be familiar to the preschool audience, but it was decided to use it anyway because it is a short word that works and sounds good in both languages.

A further adjustment was made replacing the channel Cartoonito with Cartoon Network. It was not a requirement that emerged from design activities but a design decision made because the preschool channel Cartoonito is only available in the UK and only via one multi-channel television service, Sky. Therefore it was replaced by

Cartoon Network, from which the Cartoonito channel was originated, available through several television services in the UK and Brazil.

The ‘favourite’ button represented by the heart icon was greyed out because at the beginning of the evaluation sessions the category would not contain any video. If participants added videos to favourites during the session, by pressing the ‘favourites’ button as the video is played, the button in the main menu would then become active and colourful.

There were then two versions of the prototype application, one in Portuguese (Figures 29 and 30) to be tested in Brazil, and one in English (Figure 31 and 32) to be tested in the United Kingdom. Both versions presented twelve categories. The only category that was different in the two versions was the children’s channel, CBeebies, which is not broadcast in Brazil. Therefore it was replaced by the Brazilian children’s channel TV Rá Tim Bum.

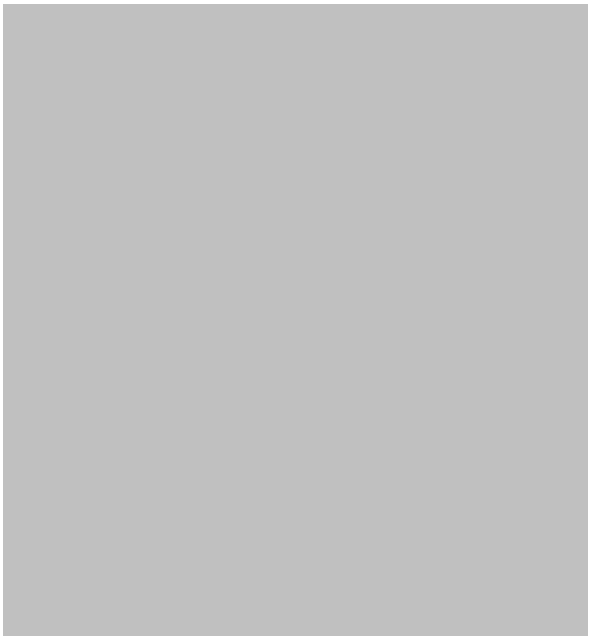


Figure 29. Portuguese version of the prototype main menu

Figure 30. Video playing on Portuguese version of the prototype

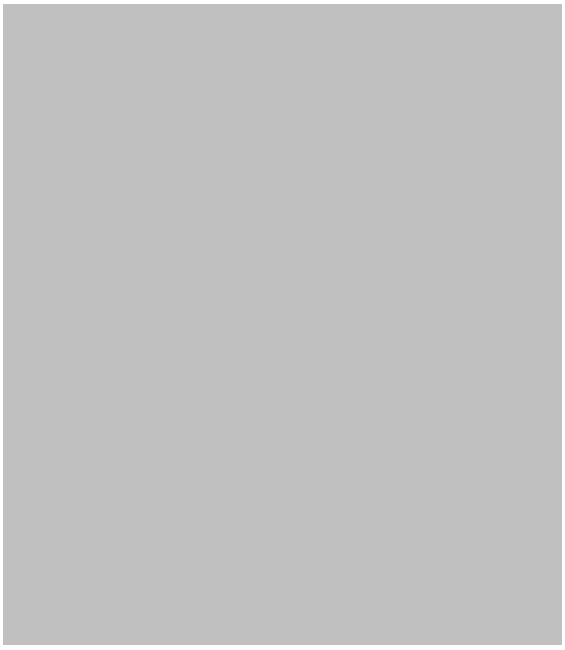


Figure 31. English version of the prototype main menu

Figure 32. Video playing on English version of the prototype

There were 83 video segments available for children to watch, so that each category had at least ten videos (RO1, RO2). All videos lasted between 30 seconds and 3 minutes. This time was chosen so that the videos would not interrupt the interactive part of the sessions for a long time if accessed accidently or even intentionally, but at the same time they would carry some meaning to enhance user experience; videos contained a song, a joke, a story segment or the programme opening.

Most videos were the same in Portuguese and English, to provide participants in both countries a similar experience, but the videos that were not found in both languages were replaced by another video of the same category. Those offered a range of regional videos that were easily accessible (REA6).

Videos presented did not take into account the time of the day (REA7) because the evaluation sessions would be carried out at different times of the day in different nurseries and it was found important to present content consistently so that children would have similar experience interacting with the prototype application.

The prototype was tested using a Dell® XPS laptop with a 15 inch widescreen display and an Intel® Core(TM)2 Duo CPU running Windows Vista®. The processor ran at 2.5GHz with 4GB of RAM and a resolution of 1280 x 800 pixels.

During the evaluation sessions, the executable prototype was used in full screen mode.

5.2.2.4 Input Devices

Participants were asked to interact with the prototype using two input devices, the mouse and the remote control. There were, however, two other alternatives to the mouse: the computer touch pad and the arrow keys and space bar in the keyboard, both options simulating the computer environment.

The mouse was the Smoby© (Figure 33) one-click USB mouse (about 9cm long and 5.5cm wide) designed for children. Instead of two buttons there is just one large button to click. This would be equivalent to the left button on a regular adult mouse and meets the requirement that all mouse buttons should have the same functionality (RL12).

Figure 33. One-click mouse

The remote control used was the One for All® (Figure 34) remote (about 19.5cm long, 5cm wide and 2cm thick), which contains the coloured buttons and arrow keys needed for interaction with the prototype (RL13). It has large buttons but the device itself is not too large for children to hold (RL1). The communication between the remote control and the computer was made via the USB infra red remote control receiver (Figure 35) and the shareware PC remote control software programmed to receive and register all remote control keys pressed. During the interaction with the remote control, the keyboard was covered with a black piece of plastic to simulate the television environment, hiding the keyboard, and to be used by participants to rest the remote if they needed (RL42).

Figure 34. Remote control

Figure 35. Infra red remote control receiver

The two alternative input devices were the touch pad (RPA1) and keyboard (RO4) built in the laptop. The interaction via these devices was provided to offer alternatives to children who may have difficulty interacting using the mouse.

5.2.2.5 Recording Software

All evaluation sessions were recorded for subsequent analysis. The 2.0 mega pixel camera built in to the computer was used to capture participants' facial expressions. Debut Video Capture© software was used to record the screen. Proprietary input monitor software recorded the time of each click and button pressed. Both webcam and screen capture software, in addition to images, recorded audio data, which facilitated synchronizing both videos to be composed into one single file. This allowed the events taking place in the screen to be viewed concurrently with participants' expressions during the data analysis process.

5.2.2.6 Adapted Fun Toolkit

The Fun Toolkit (Read and MacFarlane, 2006) survey method was adapted to be used with young children in order to gather their opinion about the input devices used during the evaluation session. The original method comprises of a form in which children are asked to tick boxes representing their answers. In the adapted version for younger children alternative materials such as stickers, crayons and rubber stamp were used.

First, participants were asked to rate each device they used on a five-point scale. They were presented with five stickers (Figure 36) and asked to choose the sticker that best represented their opinion about each device and paste it below the device image.

Figure 36. Smileyometer stickers

The second tool used was the Again-Again table, in which participants were asked which device they would like to use again. The words underneath each device image were "yes", "maybe" and "no", big and thick enough for them to colour their answer.

The third tool used was the ranking system in which children were presented with the images of the two devices and asked to use a rubber stamp with a star shape to vote for the device they thought was the best.

There were two versions of the survey form in which devices images were presented in different order. In the first version, used in conditions one and three, the mouse was presented on the left and remote on the right on both Smileyometer and Ranking sheets, and the mouse was presented on the right and remote on the left on the Again-Again table. In second version, used in conditions 2 and 4, the mouse was presented on the right and remote on the left on both Smileyometer and Ranking sheets, and the mouse was presented on the left and remote on the right on the Again-Again table. See the first version of the survey form in Appendix M.

5.2.3 Procedure

Participants were brought to the Nursery room in which the research was conducted by a member of the nursery staff. The room was set up with a table and two chairs of appropriate height for children. Participants sat on the chair in front of the computer and the researcher sat on their left. The evaluation sessions were conducted with each child individually (RL4, RL40).

The researcher introduced herself, briefly explained the research and asked if the child would like to take part. Participants were then told that they could withdraw at any time if they wished.

In order to gather some more data about their media use, participants were asked if they watched television and how they chose what to watch. Following this question they were asked if they used the computer and what they used it for.

Participants were then told they would be shown how the application works and asked to find some programmes. To avoid frustration the researcher mentioned it was a working prototype so there were only small segments of the programmes available to watch.

Participants were asked to complete six tasks, three with each input device, the order of the input devices to be used was counterbalanced, half the participants using the mouse first and half using the remote first. Since each child used both devices, two different but structurally similar sets of tasks were created and the order was randomized. As a result, there were in total four conditions (Appendix N). For all conditions, the aim of the first task was to select a video that was already on display on the 3D wheel. The second task involved selecting one category ('fairy tales' or 'superheroes') to find the video (simple search). The children were, for example, asked to find the *Cinderella* video. To do so they had to select the fairy tales category, and find *Cinderella* amongst the filtered results. Finally, in order to accomplish the third task, participants had to create a Boolean search selecting two categories to find the video ('TV shows' and 'music and songs' or 'movies and films' and 'animals and nature'). They were asked, for instance, to access a music video of the TV programme *Tweenies*, but for the *Tweenies'* screenshot to appear on the 3D wheel children had first to select both 'music and songs' and 'TV shows' categories.

Before participants were asked to complete the three tasks using a device, the tutorial for that device was shown, demonstrating how the prototype worked (RE4). The Portuguese and English versions of tutorials for both mouse and remote lasted between one minute and fifteen and one minute and twenty five seconds. The tutorials were recorded instructions composed of audio explanations and they mimicked interactions with the prototype (Figure 37). They included age appropriate instruction (RL34) on how to select a video to watch, how to return to the menu, how to select and deselect one or more categories, the fact that the greyed buttons were inactive, the option to use the ‘clear all’ button to deselect all buttons and how to access the help section again. It was meant to provide scaffold and guidance (RL9) so that no training would be needed (RL31) for children to use the system without adult assistance (RL30).

Figure 37. Excerpt from the English version of the tutorial

The evaluation session was structured so that participants watched the tutorial for the first device to be used, then were asked to find the first three videos, watched the tutorial for the second device to be used and were asked to find the three last videos.

Subsequent to the last timed task accomplished during the evaluation session, as participants watched the video they found, they were asked if they thought they could pause the video if they wanted. This question was asked in order to verify if the pause button presented on the right bottom corner of the screen was noticed.

After the tasks, participants were told they could select a video of their own choice to watch if they wished, and they were given the option to choose which device they would like to do this with.

At the end of the evaluation session, participants were asked to complete the Fun Toolkit survey (see Section 5.2.2.6). They were then complimented for helping with the prototype evaluation (RL10), thanked for their participation and given a certificate and sticker. In total each evaluation session lasted for approximately 30 minutes (RL7). A script was memorized for the sessions to be carried out consistently with all participants, see Appendix O for the detailed script.

5.2.3.1 Pilot Study

Before conducting the evaluation sessions a pilot study was carried out with two children at the One World Nursery in Brighton. This pilot study was conducted before the evaluation sessions in order to verify if the number of tasks and length of the session was appropriate and also to test the technology and materials, such as the recording software and adapted toolkit.

During this session it was noted that the screen position of the cursor varied each time the prototype menu was loaded. It took participants some time to find the cursor. Hence the prototype was amended so that every time the main menu was loaded the cursor would appear consistently in the middle of the screen (RL20).

One of the tasks participants were asked to accomplish involved finding the film *Ratatouille,* a Disney animation movie about a young rat who dreams of becoming a French chef. The first button one of the participants in the pilot test selected in order to find the film was ‘animals and nature’. However, the ‘animals and nature’ category in the prototype included only real animals. According to child psychology literature children in the preoperational stage may confuse fact with fiction (Meggitt, 2007) and may also believe that characters in cartoons are real (Shaffer, 1999). Therefore, following the pilot test it was decided to include in the ‘animals and nature’ category not only real animals but also animated movies and cartons in which one or more animals were the main characters.

In the remote control mode it was noticed during the pilot test that the 3D wheel displaying the screenshots stopped moving when children moved the cursor to select the buttons. As a result participants could not see all the programmes from the button selected on screen unless they navigated back to the wheel and did it manually by pressing the arrow keys on the remote control. This was a bug in the prototype, which was fixed for subsequent evaluation sessions so that the 3D wheel would move continuously in both mouse and remote control modes.

The pilot study resulted in the prototype adjustments mentioned above and it was also useful to assure the appropriateness of the tasks and the structure of the evaluation session. Participants were familiar with most videos they were asked to find, so these were maintained. It was found however that children would need some hints in order to accomplish the six tasks in 30 minutes and such assistance would also provide them a better experience, preventing participants getting stuck. Thus, it was decided to assist participants when necessary and during the data analysis to reflect upon this assistance, instead of asking participants to interact on their own with no interference, which could result in fewer tasks accomplished and a more stressful session.

5.2.3.2 Hints

The concept of scaffolding and the importance of its use in technology for children (RL9) had been emphasized in Chapter 4. Following the pilot test it was decided that scaffold and guidance should be a requirement for the design of evaluation sessions as well and could be presented in the form of hints.

Hutchinson (2005) had made use of hints in her work in order to avoid participants struggling to complete the tasks and exceeding the 30 minute time limit recommended for sessions with young children. She highlighted the problems of inconsistency and bias that could be caused by using hints and suggested using a protocol for giving hints in a consistent way.

Following Hutchinson's recommendations and in line with the problems encountered in the pilot test, a list of hints was created for tasks using both devices and reported during the analysis.

There were five types of hints:

- Hints about mouse navigation
- Hints about remote control navigation
- Hints to look at the buttons on the prototype menu
- Hints about a category
- Hints about alternative input devices

For the first issue, regarding mouse navigation, if participants needed help in selecting the video to watch or in selecting/deselecting a category they were asked if they remembered how to do so. If they didn't they were told to click (on the screenshot if they wanted the video to appear on the whole screen, or on the button to select/deselect a category). Where participants did not understand they had to aim at the elements on the screen before clicking, they were given an additional hint (Hint 2). Participants were also given hints about the shortcuts (to use the 'clear all' and 'back' buttons) if they needed.

Hint 1 – You should click

Hint 2 – You should move the little hand (pointer) so it is over it

Hint 3 – You could click on the yellow button to deselect all buttons

Hint 4 – You could click on the red button to go back

In order to assist with the interaction via the remote a similar procedure was used. Firstly participants were asked if they remembered how to select the video to watch on the whole screen or select/deselect a category. When they did not recall they were told to press the OK button. If necessary, participants were also given hints on how to navigate with the remote, using the arrow keys and the coloured buttons.

Hint 5 – You should press OK

Hint 6 – You should use the arrow buttons to move the little hand (pointer) so it is over it

Hint 7 – You could use the yellow button to deselect all buttons

Hint 8 – You could use the red button to go back

The revolving screenshots were found sometimes to distract participants and limited their attention to the upper screen of the prototype. It was found necessary to have a hint to direct them to look at the buttons on the lower part of the application for the tasks in which they had to select categories.

Hint 9 – Take a look on the buttons on the bottom of the screen

Participants were also given a hint when they needed assistance to select the appropriate categories to find a video.

Hint 10 – For simple searches – We are looking for (name of the programme). It is a (category name). Where do you think we can find (category name)?
Hint 11 – For Boolean searches – We are looking for (name of the programme). It is a (category 1 name and category 2 name). Where do you think we can find (category 1)? How about (category 2)?

If the wrong category was selected, participants were told it was not right (Hint 12) and then given Hints 10 or 11 depending on the type of the search being conducted.

Hint 12 – That's not right, is it?

Hints were also provided in case participants wanted or needed to use an alternative input device. There were two hints regarding the touch pad navigation.

Hint 13 – You should press the left button on the touch pad
Hint 14 – You should move the little hand (pointer) so it is over it

For issues with the keyboard navigation there were also two hints.

Hint 15 – You should press the spacebar
Hint 16 – You should use the arrow keys to move the little hand (pointer) so it is over it

5.3 Results

The results were analysed as soon as all the evaluation sessions have been conducted. A professional statistician was consulted to ensure the quantitative data was analysed properly. The tests used, briefly described below, are based on the combination of data collected through webcam, screen capture and input monitor software. Most data gathered was not normally distributed, so non-parametric statistics techniques were used.

For all statistics the null hypothesis tested was that there is no difference between the groups of data being compared, the cut-off probability value used (p value) was 0.05 to reject the null hypothesis in favour of an alternative hypothesis that the two groups of data differ. The small number of participants, however, may have not been enough to detect differences present in the data. For this reason, the data analysed statistically was also illustrated with graphs, tables and, when appropriate, compared with the results from the video analysis. Most graphs are illustrated using boxplots to provide graphical display of the local and spread of variables. In the boxplots the line in the box indicates the median value of the data. The ends of the vertical lines are the minimum and maximum data values. When outliners are present they are points presented outside the box and lines. Two additional pieces of information were added to the boxplots to facilitate the exploratory data analysis. These are the mean represented by a cross inside a circle, and the mean connect line which connects the means of the samples being analysed.

For the qualitative data analysis the video data was transcribed and coded. Following the transcription and coding for all videos the codes were reviewed and organized into categories. Some categories derived quantified data such as participants' responses about their own media use, and the different types of hints given to assist participants accomplishing tasks. Other categories elucidated details of participants' understanding of the prototype application and the interaction process.

The variables examined were: tasks accomplished, time taken to accomplish tasks, hints given to accomplish tasks, interactions to accomplish tasks, participants' interactions with the prototype and participants' preferences for input device.

The statistical calculations and most graphs were produced using the statistics package Minitab 15 while the qualitative data analysis software NVivo 8 was used to transcribe and analyse the video data. In this chapter, participants' quotes are in italic and the quotes taken from participants in Brazil are translated into English.

5.3.1 Measuring Interactions with the Prototype

5.3.1.1 Tasks Accomplished

Participants were asked to accomplish six tasks in total, three with each device. The video data from the evaluation sessions was then analysed to verify the tasks participants successfully accomplished. Some tasks were not accomplished because participants accessed a video other than the one they were required to, other tasks were interrupted because participants were found to be tired or fidgety in struggling to accomplish it. A series of tests were used to analyse if there was a difference in the number of tasks accomplished by; participants' age, gender, the country context, the condition they were submitted (one, two, three or four), participants' media and device usage, input device used, and task type. These results are presented in Section 5.3.2.

5.3.1.2 Time Taken to Accomplish Tasks

It was also investigated and it is presented in Section 5.3.2, the amount of time in seconds taken to accomplish tasks by participants and the factors that could have influenced the duration of each task.

5.3.1.3 Hints Given to Accomplish Tasks

As mentioned previously in this chapter (Section 5.2.3.2) hints were given to participants in case they were found to be stuck, puzzled, or were taking a long time to accomplish tasks. Hints were divided into five types: there were hints about mouse navigation, hints about remote control navigation, hints to look at the buttons on the prototype menu, hints about categories and hints about alternative input devices. A large number of hints of all types were given to participants during the evaluation sessions, but among the results are only considered hints given to participants during tasks that were actually accomplished.

A total of 195 hints were given to assist participants to accomplish tasks, but these hints are divided into 16 hints of 5 different types. Thus, there is not enough data to provide evidence that among the distinct hints the number of times each hint was given is statistically significant different. However, it may be noticed that the hints that were most given were hints five and six, which regard remote navigation, and hint eleven that indicates the two categories in a Boolean search (Figure 38).

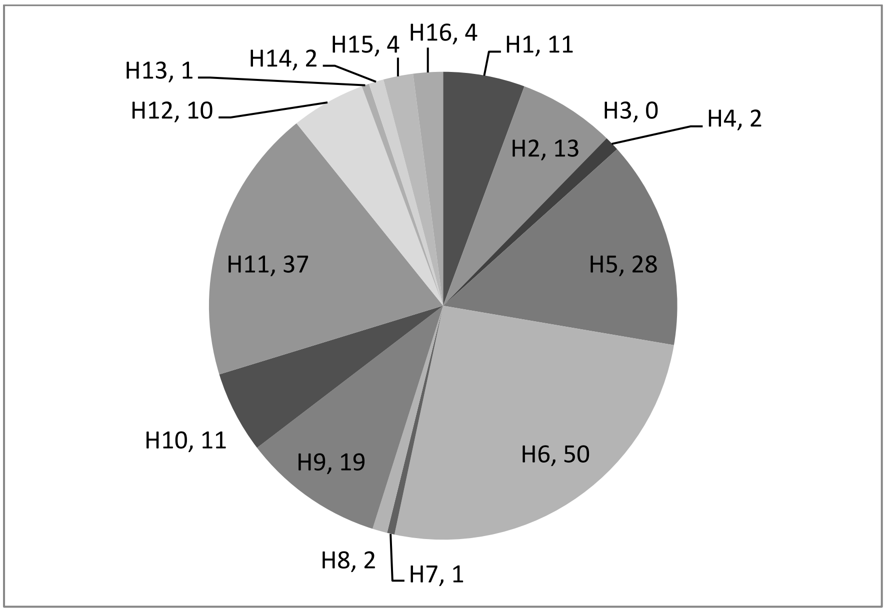


Figure 38. Number of hints given divided into each individual hint

It can be seen that a larger number of hints were given to assist participants with the remote navigation than to help participants navigate using the mouse (Figure 39).

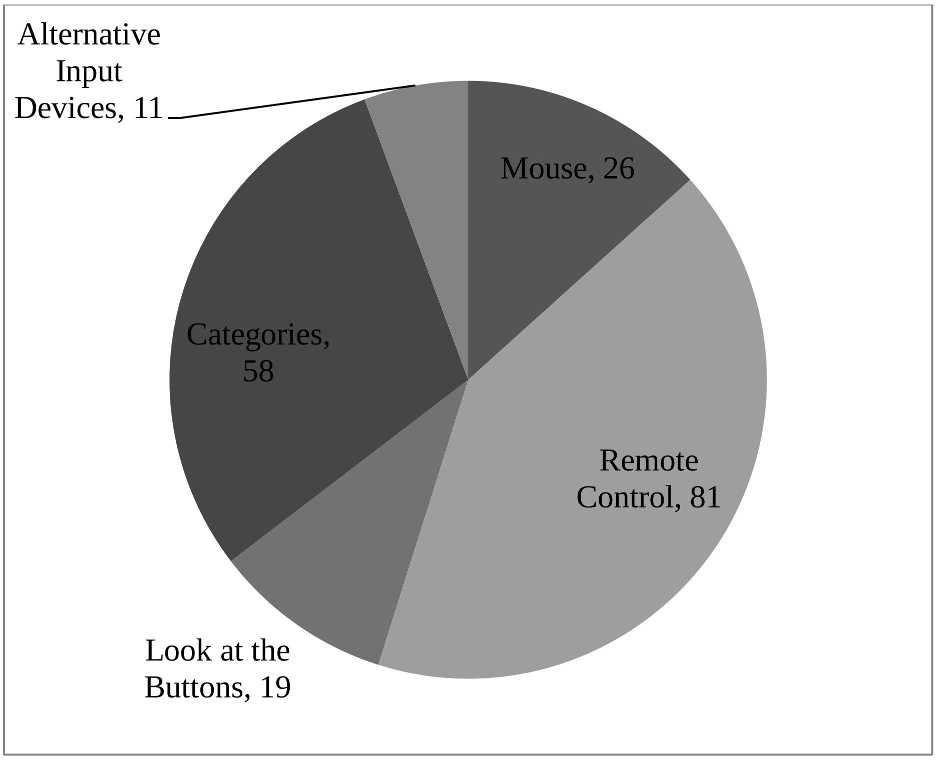


Figure 39. Number of hints given divided into hint types

The number of hints given according to participants' age, gender, country context and condition will be presented and discussed on the following sections along with a further discussion on the hints given by participants to interact with each input device and also according to the task types and participants' media and device usage.

5.3.1.4 Interactions to Accomplish Tasks

During the evaluation sessions, the input monitor software recorded the time of each click and button pressed. Then, at the data analysis stage, it was calculated the number of clicks and buttons pressed to accomplish each task in order to verify how many interactions were needed to accomplish tasks. In the following section it is analysed if there was a difference on the number of interactions for each task accomplished according to participants' age, gender, country context and condition, as well as participants' media and device usage, the input device used, and type of task accomplished.

5.3.2 Factors that Could Have Affected Interactions with the Prototype

At first, non-parametric statistical methods were used to calculate the factors affecting children's interactions with the iTV prototype. As previously mentioned, the null hypothesis tested was that there is no difference between the groups of data being compared (e.g. two genders, two country contexts, different conditions), the cut-off probability value used (p value) was 0.05 to reject the null hypothesis in favour of an alternative hypothesis that the two groups of data differ. The resulted p values are presented in the tables 2 and 3 below and further discussed in the following subsections.

	Tasks Accomplished	Time to Accomplish Tasks	Hints Needed	Interactions Performed
Age	p value=0.03	p value=0.1104	p value=0.2623	p value=0.9737
Gender	p value=0.8691	p value=0.4343	p value=0.9212	p value=0.3068
Country Context	p value=0.0078	p value=0.4097	p value=0.0256	p value=0.0878
Condition	p value=0.739	p value=0.126	p value=0.930	p value=0.920

Table 2. Resulted p values for the different variables examined

Media and Device Use	Tasks Accomplished	Time to Accomplish Tasks	Hints Needed	Interactions Performed
TV	p-value=0.03	p value=0.441	p value=0.016	p value=0.057
Computer	p value=0.269	p value=0.121	p value=0.207	p value=0.127
Remote	p value=0.272	p value=0.289	p value=0.101	p value=0.121
Mouse	p value=0.297	p value=0.801	p value=0.253	p value=0.024

Table 3. Resulted p values by participants' media and device use

The results in Table 2 indicate that age and the country context affected children's interactions with the prototype. The p values of tasks accomplished by participants' age and country context are smaller than 0.05 indicating that the number of tasks accomplished by participants of different ages and country contexts is different. The p value of hints needed to accomplish tasks by country context is also smaller than 0.05, the null hypothesis should then be discarded in this case as well, the number of hints given to participants in Brazil was different than the number of hints given to participants in the UK.

The p values presented in Table 3 indicate that the frequency children watch television may affect the number of tasks they accomplished during the evaluation sessions and the hints they needed to accomplish tasks. The p value for interactions performed according to children's mouse usage is also smaller than 0.05, indicating that the number of interactions performed to accomplish tasks is different for participants who use the mouse unsupervised, those who use the device only with supervision and participants who do not use the mouse.

All other p values are greater than 0.05, but this not necessarily means that the null hypothesis is true, it only suggests that there is not enough evidence in favour of the alternative hypothesis. The small sample, as already mentioned, may have not been enough to detect differences present in the data. Thus, to investigate which factors actually interfered with children's interactions with the iTV prototype during evaluation sessions, each factor is analysed further in the following subsections. In addition to the participants' characteristics mentioned above it is also investigated how the input device and the task type affected the interaction with the prototype.

5.3.2.1 Age

An analysis of correlation was performed on the dependent variable, number of tasks accomplished, for participant's age in months. It was found that the correlation between age and accomplished tasks is significant (p value=0.03), older participants accomplished more tasks than the younger participants (Figure 40).

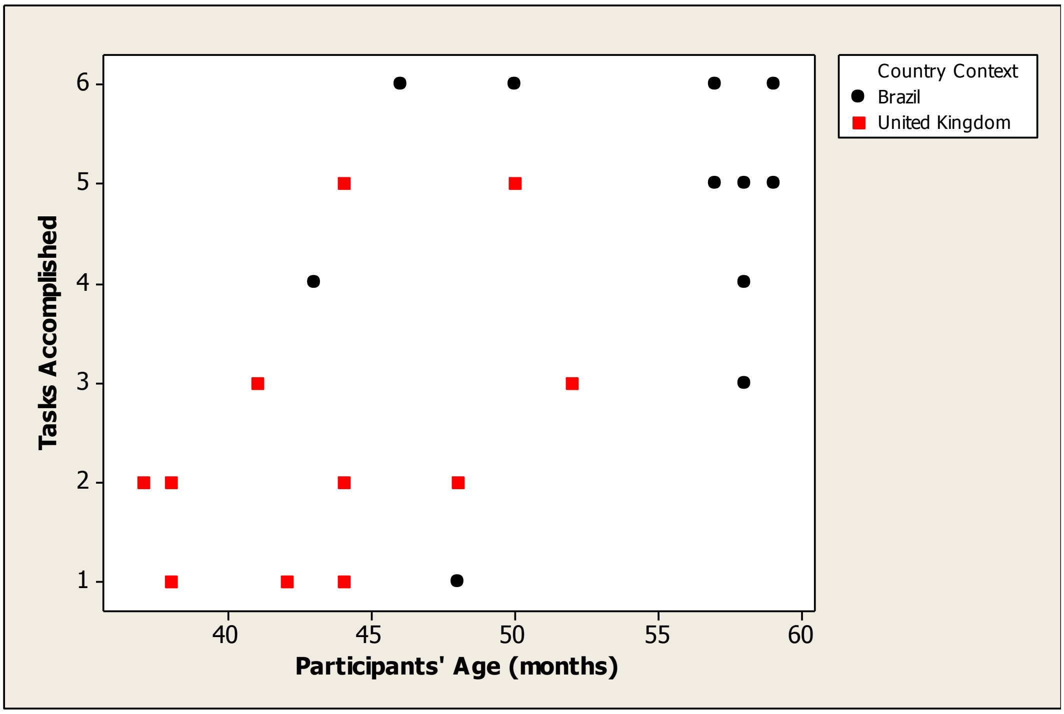


Figure 40. Number of tasks accomplished by participants' age

A Mann-Whitney test was used to calculate the effect that age had on the time on average that participants took to accomplish the tasks. The result (p = 0.1104) suggested that the time taken by three years old participants to accomplish tasks is not statistically different than the time taken by four years old participants. It can be noticed, however, from the boxblot below (Figure 41) that three year olds took slightly longer to accomplish tasks than four year olds and there was more variation on the time three years old took to accomplish tasks.

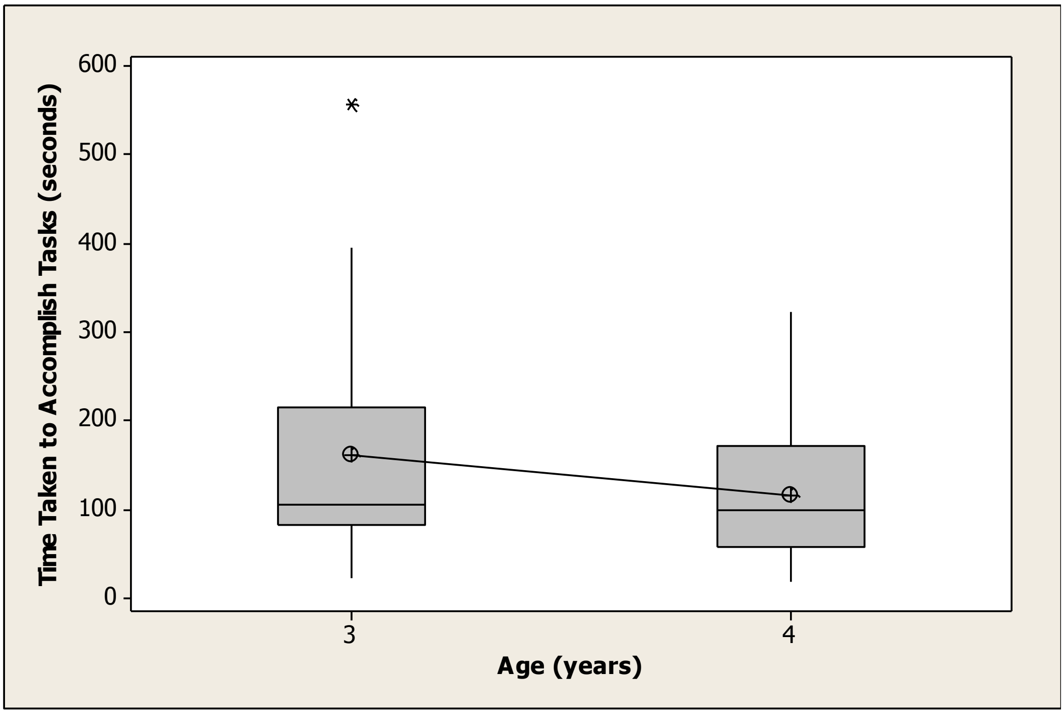


Figure 41. Average time taken to accomplish tasks by participants' age

Another Mann-Whitney test was carried out and indicates that the number of hints given to three and four years old participants is not statistically significant different (p value = 0.2623). On average, however, both mean and median, older participants were given more hints than younger participants (Figure 42).

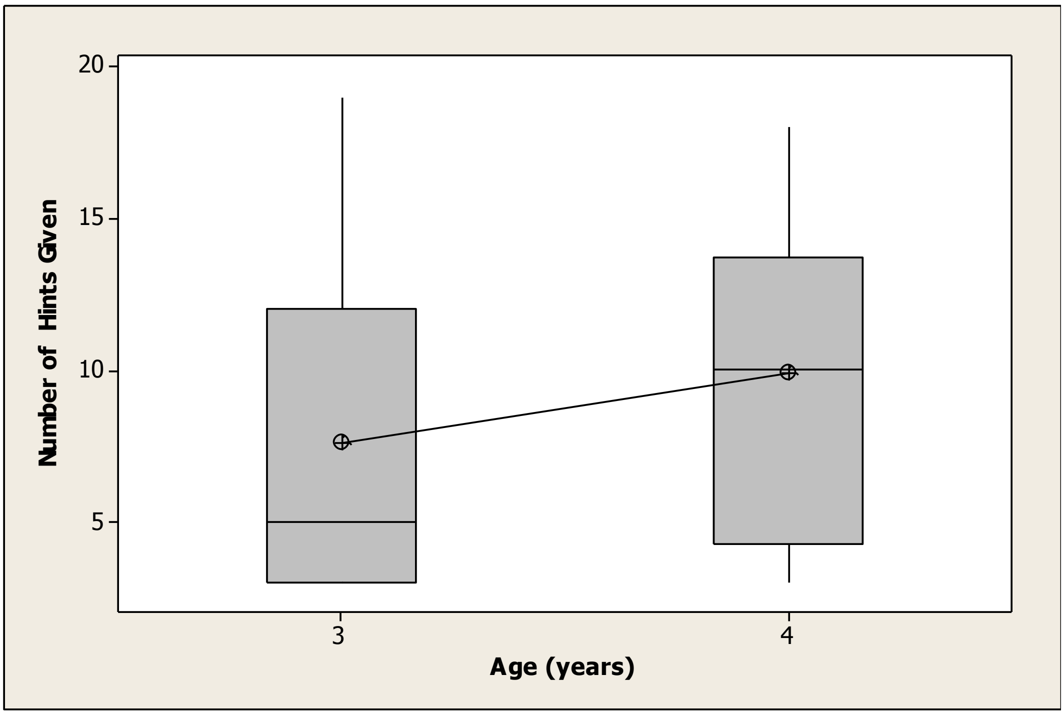


Figure 42. Average number of hints given by participants' age

Those are, however, the total amount of hints, including hints from all types. If divided into the different types, there is not enough evidence that there is a statistically significant difference on the average number of hints to look at the buttons on the bottom of the screen (H9) given to three and four years old participants (p value = 0.366). There is not enough evidence either that there is a statistically significant difference on the average number of hints about categories (H10, H11 and H12) given to participants on the two age groups (p value = 0.212). Hints regarding the input device navigation (H1, H2, H3, H4, H5, H6, H7, H8, H13, H14, H15 and H16) will be discussed in Section 5.3.2.6.

An additional Mann-Whitney test was carried out and the results suggest that there is no statistically significant difference on the number of interactions, clicks and buttons pressed, performed to accomplish tasks for three and four years old participants (p value = 0.9737). It may, however, be noticed that on average, both mean and median, three year olds needed a slightly larger number of interactions to accomplish tasks compared to four year olds (Figure 43). On average (mean) three years old participants performed 68.8 interactions to accomplish tasks while four years old participants performed 55.1.

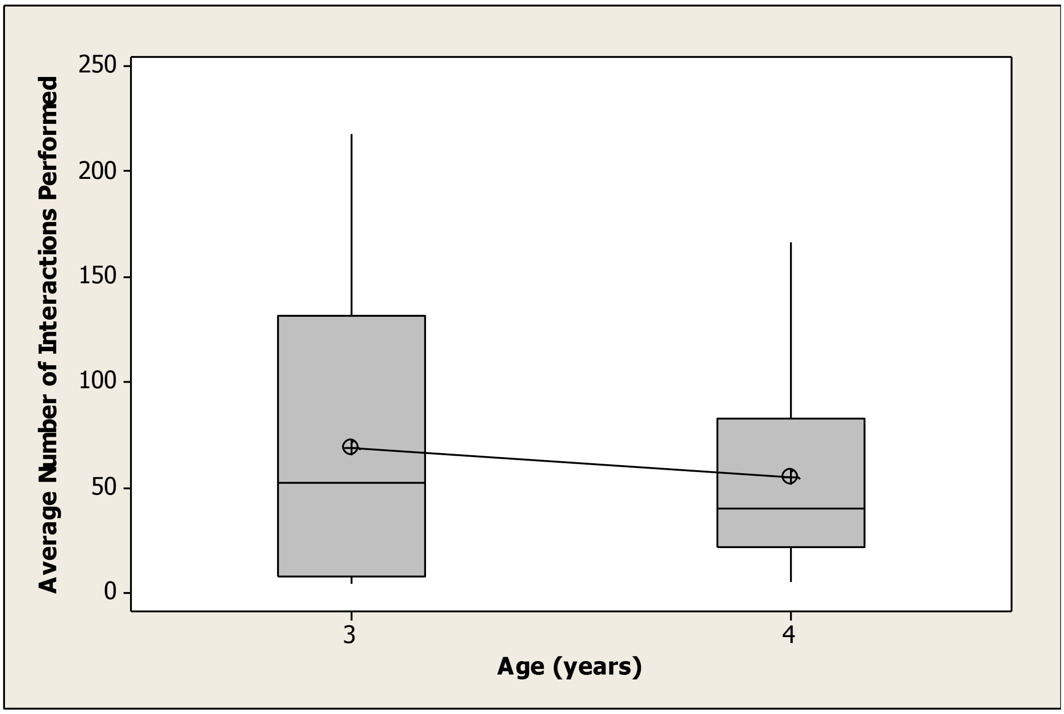


Figure 43. Average number of interactions performed by participants' age

5.3.2.2 Gender

During the video analysis it was noted that there was no evident difference between the performance of boys and girls in evaluation sessions conducted in the UK. In Brazil, however, boys appeared to be more confident interacting with the prototype while girls were to some extent doubtful and shy. Participants' quotes presented below illustrate such difference in behaviours among participants in Brazil.

Boys in Brazil talked about what they knew and how easy it was to accomplish tasks: *I know everything* (Participant 1). *I know how this works* (Participant 2). Participant

6, also a boy, mentioned six times during the session tasks were easy to accomplish, before and after accomplishing the tasks.

Girls in Brazil, on the other hand, not only did not make such statements but also showed signs of lack of confidence: Participant 4 asked the researcher seven times during the session *what do you mean?* Referring to the questions at the beginning of the session, to the tasks to be accomplished and also to the stickers she could choose as a reward at the end of the session. Participant 8 did not speak a single word during the entire 30 minute session. Participant 10 also demonstrated her lack of confidence, despite the fact that she successfully accomplished the six tasks, during the second to last task she told the researcher *I can't do it*.

In order to see if there was a difference in the number of tasks accomplished by gender a Mann-Whitney test was carried out and the result (p value=0.8691) indicates that the median of tasks accomplished by girls is not significantly different to the median of tasks accomplished by boys. Another Mann-Whitney test was conducted including only participants in Brazil, to verify if the difference on behaviour among the genders noticed through video analysis affected the number of tasks participants accomplished. Results indicate that there is no significant difference (p value = 0.3613) between the number of tasks boys and girls accomplished during the sessions in Brazil.

There was no statistically significant difference in the number of tasks girls and boys had accomplished. However, it was found that there was a difference in behaviour among participants during the evaluation sessions in Brazil; boys appear to be more confident interacting with the prototype than girls. In order to analyze if such difference in behaviour affected the time participants had taken to accomplish tasks a Mann-Whitney test was carried out. Test results indicate that there is no significant difference (p value = 0.4343) on the time girls and boys had taken to accomplish the tasks. Another Mann-Whitney test was conducted with only results from participants in Brazil and it also resulted in no significant difference (p value = 0.8176) on the time boys and girls took to accomplish tasks. In fact the graph below shows (Figure 44) that girls in Brazil had accomplished tasks in average, both mean and median, slightly faster than boys.

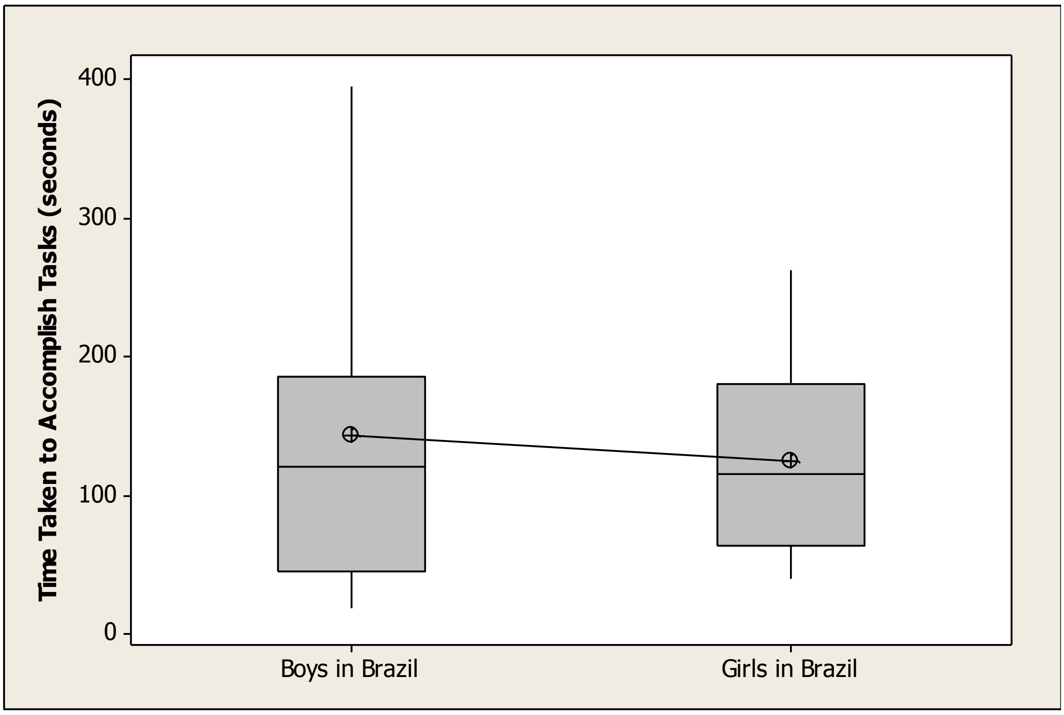


Figure 44. Average time to accomplish tasks by boys and girls in Brazil

Another Mann-Whitney test was conducted and the result suggests that there is no statistically significant difference on the number of hints given to boys and girls (p value = 0.9212). In this case, it can be observed from the boxplot below (Figure 45) that the average, both mean and median, number of hints given to boys and girls is very similar.

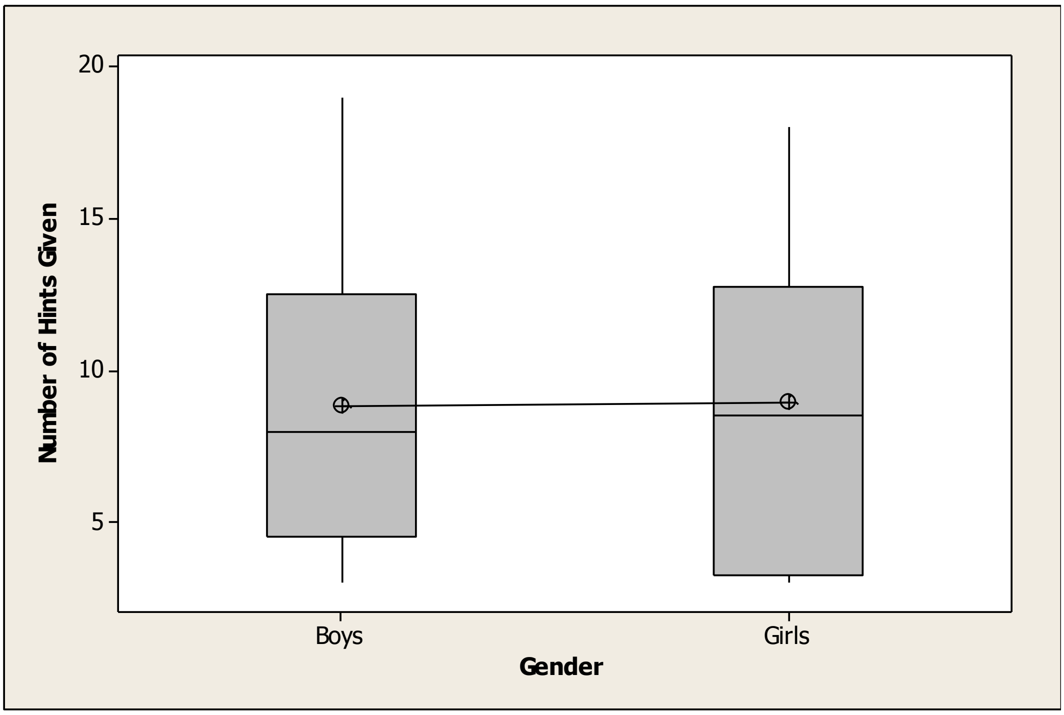


Figure 45. Average number of hints given by participants' gender

According to the different hint types, there is not enough evidence that there is a statistically significant difference on the average number of hints to look at the buttons on the bottom of the screen (H9) given to boys and girls (p value =0.439). There is not enough evidence either that there is a statistically significant difference on the average number of hints about categories (H10, H11 and H12) given to participants of both genders (p value = 0.982).

A Mann-Whitney test was conducted and suggests that there is no statistically significant difference on the number of interactions needed to accomplish tasks by boys and girls (p value = 0.3068). The average number of interactions performed by participants of both genders was similar.

5.3.2.3 Country Context

In this thesis, country context refers to the setting of the study, i.e. the countries the evaluation sessions were held in, Brazil and the United Kingdom, which were probably participants' countries of residence. It is worth noting that environmental, ethnic, cultural and social, among other variations, may occur between and within countries.

Based on the country context, a Mann-Whitney test was carried out and the result (p value=0.0078) indicates that participants in Brazil accomplished between one and four tasks more than participants in the UK. However the mean age of participants in Brazil was 54 months while the mean age of participants in the UK was 44 months, participants in Brazil were significantly older than those in the UK; therefore there is not enough evidence that the number of tasks participants accomplished was associated with the country context.

During the video analysis, however, there is evidence that participants in Brazil and in the UK behave differently in the evaluation sessions. Participants in Brazil did not question the task to be accomplished by, for example, asking to access a video other than the one they were required to find. Only one participant in Brazil while accomplishing a task said *I want to watch films* (Participant 11), but it is not clear if he meant he wanted to access a film instead of the video he was asked to find or if he was referring to films in general that could include the video to be found. Several participants in the UK, on the other hand, asked to watch a video of their choice instead of the videos part of the task. Participant 12 was asked to find the film *Cinderella* then told the researcher *I want to watch Charlie and Lola*. Participant 13 was asked to find the *Dora the Explorer* video she closed the prototype when told to open it to find *Dora* she said *no, I want to watch something else, I want to watch a cartoon*. Participant 17 when asked to find the *Dora the Explorer* video also told the researcher *I want to watch Charlie and Lola.* Participant 18 was asked to find the animation *Pingu* but said *I want to see Lazy Town* and later in the same evaluation session when asked to find *Cinderella* the participant said *I want this one now,* pointing at the screenshot from the *Charlie and Lola* cartoon displayed on the screen. Participants in the UK also showed more signs of fidgety and tiredness during the evaluation sessions than participants in Brazil. Only one participant in Brazil (Participant 8) showed signs of fidgety and tiredness while eight out of eleven participants in the UK looked at some point during the evaluation session tired or fidgety.

To compare the time taken by participants of both country contexts to accomplish tasks a Mann-Whitney test was carried out and suggested that there is no significant difference (p value = 0.4097) between the time participants in Brazil and in the UK took to accomplish the tasks.

In order to measure the hints given according to country context a Mann-Whitney test was carried out and the result indicates that there is a statistically significant difference on the number of hints given to participants in Brazil and in the UK (p value = 0.0256). Participants in Brazil were given on average more hints than participants in the UK (Figure 46).

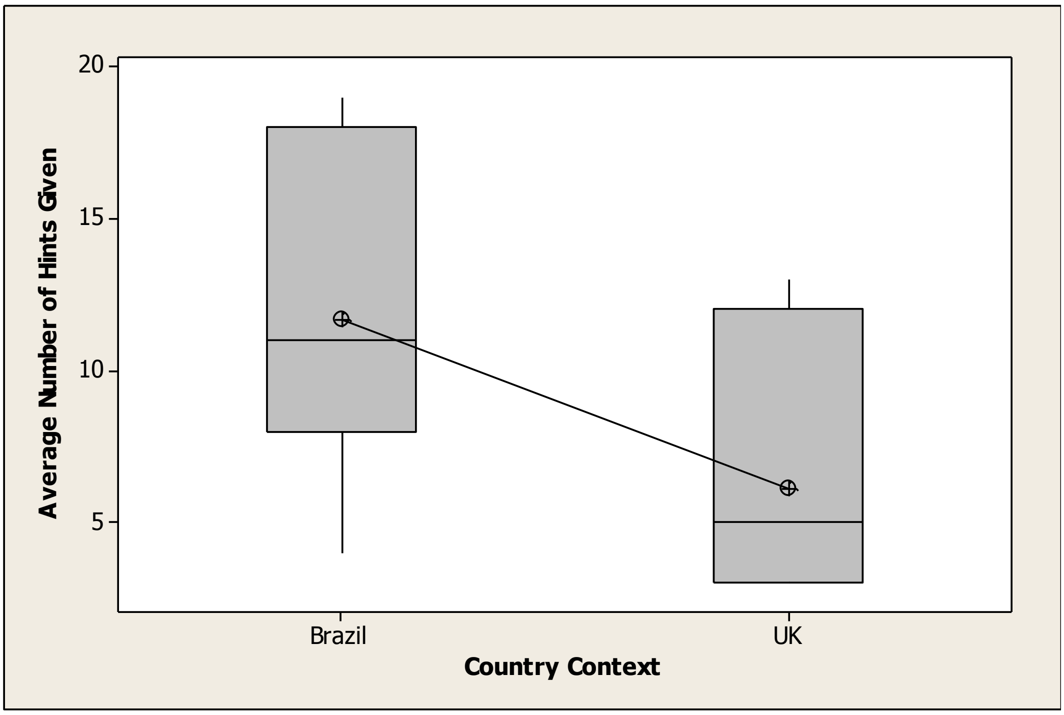


Figure 46. Average number of hints given by country context

Regarding the different hint types, there is not enough evidence that a statistically significant different amount of hints to look at the buttons on the bottom of the screen (H9) was given for participants in Brazil and in the UK (p value = 0.111). According to the hints about categories (H10, H11 and H12) there is not enough evidence either that there is a statistically significant difference on the amount of hints given to participants in Brazil and in the UK (p value = 0.088). As previously mentioned, hints that regard input devices are reported later.

Another Mann-Whitney indicates that there is not enough evidence that the number of interactions needed to accomplish tasks by participants in Brazil and in the UK is statistically significant different (p value = 0.0878). Participants in Brazil, however, performed on average, mean and median, more interactions to accomplish tasks than participants in the UK (Figure 47).

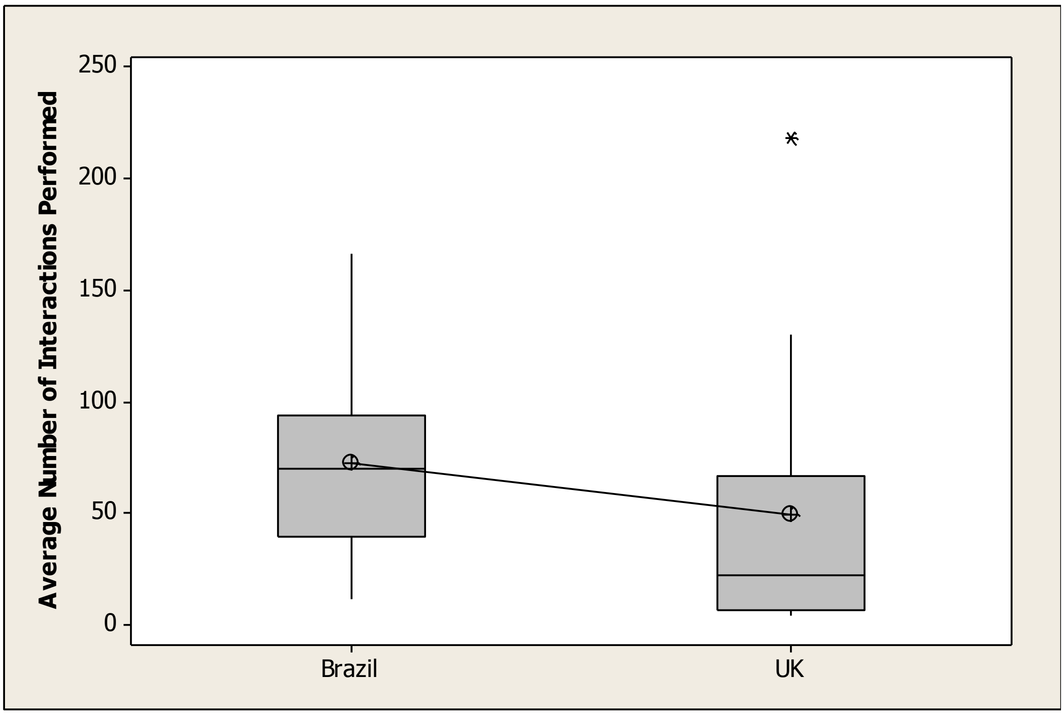


Figure 47. Average number of interactions performed by country context

5.3.2.4 Condition

As previously mentioned there were in total four conditions. In condition one, participants were asked to accomplish the first set of tasks, using the mouse first and the evaluation form one. In condition two, participants were asked to accomplish the first set of tasks, using the remote first and the evaluation form two. In condition three, participants were asked to accomplish the second set of tasks, using the mouse first and the evaluation form one. In condition four, participants were asked to accomplish the second set of tasks, using the remote fist and the evaluation form two.

A Kruskal-Wallis test indicated that the condition did not affect the results (p value=0.739), therefore the null hypothesis was not discarded. The number of tasks accomplished by participants in conditions one, two, three and four was not statistically significantly different.

Another Kruskal-Wallis test was performed to verify if the condition participants were submitted affected the time they accomplished the tasks. It was found that there was no significant difference (p value = 0.126) on the average time participants took to accomplish the tasks on the four different conditions.

There is no statistically significant difference on the number of hints given to accomplish tasks for each one of the four conditions (p value = 0.930). Looking at the different types of hints individually, a Kruskal-Wallis test indicates that there is no evidence of a statistical difference on the numbers of hints to look at the buttons on the bottom of the screen (H9) for participants on the four different conditions (p value = 0.273). There is no evidence of a statistically significant difference either on the number of hints about categories (H10, H11 and H12) given to participants on the four different conditions (p value = 0.182).

An additional Kruskal-Wallis test indicates that there is no statistically significant difference on the average number of interactions performed to accomplish tasks for each one of the four conditions participants were submitted to (p value = 0.920).

5.3.2.5 Participants' Media and Device Use

In this section first it is presented an overview of the findings from the data on participants' media and device usage gathered through the parents' questionnaires and participants responses about their own media use. It is then calculated if there is a difference on the number of tasks accomplished according to participants' media and device usage. Please note that the tables illustrating media use only include the data that is different from zero for at least one of the samples being compared. For frequency that participants watch TV, for instance, none of the parents responded "Never" on the questionnaire, so "Never" was not included on the table illustrating the frequency participants watch TV.

5.3.2.5.1 Overview of Participants' Media and Device Usage

The data presented below reflects the responses from the parents' questionnaire about children's' media and device usage according to participants' age, gender and country context. There is not enough data to generate statistically relevant results so the numbers are illustrated in tables.

There is not enough evidence that the frequency of participants' media use is different among the two age groups, but it may be noticed that older participants on average watch television and use the computer more frequently than younger participants (Tables 4 and 5).

Frequency participants watch TV	Number of 3 years old participants	Number of 4 years old participants
Several times per day	3	7
About once a day	4	5
A couple of times a week	3	0

Table 4. Frequency participants watch TV according to their age

Frequency participants use the computer	Number of 3 years old participants	Number of 4 years old participants
Several times per day	0	1
About once a day	2	3
A couple of times a week	3	2
About once a week	1	3
About a couple of times a month	1	1

Frequency participants use the computer	Number of 3 years old participants	Number of 4 years old participants
About once a month	1	0
Less often	2	0
Never	1	0

Table 5. Frequency participants use the computer according to their age

Regarding the device usage, there is not enough data either to provide evidence that the mouse usage of three years old participants and four years old participants is different, but on average a larger number of older participants make use of the devices with no supervision compared to younger participants. Four out of ten three years old participants use the remote without being supervised while nine out of the twelve four years old participants use the remote with no supervision. According to the mouse usage, eight out of the ten three year olds use the mouse without being supervised while eleven out of the twelve four year olds use the mouse with no supervision.

Among the two age groups the remote control usage varies slightly more than the mouse usage. A somewhat larger number of older participants use the remote with no supervision compared to younger participants while the number of participants who use the mouse is similar for the two ages. There is not enough evidence, however, that participants' age combined with their device usage had an effect on the number of tasks accomplished. A larger number of three years old participants use the mouse without supervision compared with the remote, for instance, but as discussed and illustrated in Section 5.3.2.6 within each age, the average number of tasks accomplished with both devices is similar.

There was not enough evidence that there was a difference on the media usage by participants' gender. Boys and girls watch TV on average with similar frequency (Table 6).

Frequency participants watch TV	Boys	Girls
Several times per day	6	4
About once a day	2	7
A couple of times a week	2	1

Table 6. Frequency participants watch TV according to their gender

Regarding the computer usage, participants on both genders also have similar patterns (Table 7).

Frequency participants use the computer	Boys	Girls
Several times per day	0	1
About once a day	1	4
A couple of times a week	3	2
About once a week	3	2
About a couple of times a month	0	2
About once a month	1	0
Less often	2	0
Never	0	1

Table 7. Frequency participants use the computer according to their gender

The device usage is also similar for both genders. For the remote a slightly higher number of boys use it without supervision compared to the girls, for the mouse the numbers of boys and girls that use it with no supervision is very similar. Seven out of the ten boys use the remote without being supervised while six out of the twelve girls use it with no supervision. Considering the mouse, nine out of the ten boys use the device with no supervision and ten out of the twelve girls use it without being supervised.

Considering participants' country context there was not enough evidence to indicate a statistically significant difference on the frequency of media use among children in Brazil and in the UK. It can, however, be noticed from the Table 8 below that on average participants in Brazil watch television more frequently than participants in the UK.

Frequency participants watch TV	Participants in Brazil	Participants in the UK
Several times per day	8	2
About once a day	3	6
A couple of times a week	0	3

Table 8. Frequency participants watch TV according to their country context

The frequency participants from both country contexts use the computer is, on the other hand, very similar (Table 9).

Frequency participants use the computer	Participants in Brazil	Participants in the UK
Several times per day	1	0
About once a day	2	3
A couple of times a week	1	4
About once a week	4	1
About a couple of times a month	1	1
About once a month	0	1
Less often	1	1
Never	1	0

Table 9. Frequency participants use the computer according to country context

Regarding the input device usage, a larger number of participants in Brazil use the remote control with no supervision compared to participants in the UK. Nine out of eleven participants in Brazil use the remote control without being supervised while only four out of eleven participants in the UK use the remote with no supervision. The mouse usage is more alike on both country contexts. Nine out of eleven participants in Brazil use the mouse without being supervised while ten out of eleven participants in the UK use the mouse with no supervision.

At the beginning of the evaluation sessions, participants were asked if they watch television and twenty one out of twenty two participants affirmed they watch television. The only participant (Participant 2) who said he does not watch TV said that he used to watch it but the television set he has in his room broke down.

Following this question participants were asked how they choose what to watch. Two participants mentioned they choose what to watch by selecting a channel. Participant 3 when asked if he watches TV he mentioned *I watch Cartoon and Discovery* then as asked how he selects what to watch he said *I press the five and the seven and four and five.* The channels 57 and 45 are equivalent to the Cartoon Network and Discovery Kids channels respectively on a Brazilian cable television provider (NET). Participant 13 when asked how she chooses what to watch said *I choose cartoons and channels and CBeebies*. Four participants said they choose a disc (CD or DVD). Participant 2 said *there are a lot of CDs there*. Participant 19 said *we've got Madagascar 2, we've got a big box that's got two CDs in it, the Christmas one and the Penguin*. Four participants choose by a specific programme, Participant 5 said she watches *Cinema em Casa* (Brazilian afternoon film session broadcast on weekdays), Participant 10 said she chooses *Cocoricó* (Brazilian television Puppet Show), Participant 11 mentioned he chooses the cartoon *Woody Woodpecker* and Participant 22 said she chooses *Peppa Pig* (an animated British TV series). Three participants have someone else to choose or assist them choosing what to watch. Participant 6 said *my sister rents a film*, Participant 1 mentioned his grandmother watches TV with him and chooses what to watch and Participant 20 said *my mommy or my daddy helps me*. The other participants did not respond the question.

Participants were then asked if they use the computer. Fifteen participants affirmed they use the computer, four participants mentioned they do not use the computer and three participants did not answer the question.

The last question was about their computer use, for those who said they use the computer it was asked what they use it for. Three participants mentioned they use the computer as a tool to paint, draw and write. Participant 2 said he uses the computer to paint, Participant 7 said she uses it to make drawings and Participant 10 said *my mom let me write sometimes on the computer*. One participant (Participant 9) mentioned he uses the computer to play games, while another participant (Participant

15) said he uses it for pastime. Two participants mentioned they visit broadcasters' websites. Participant 6 said *I use the Discovery's website*; Participant 20 said *I sometimes do some CBeebies' or something.* The other participants did not answer the question.

Participants' responses about their own media use were interesting, but a large number of participants, especially the younger ones, did not answer one or more questions they were asked, as a result, all calculations involving participants' media and device usage were based only on the data gathered through parents' questionnaires.

5.3.2.5.2 Interactions According to Media and Device Usage

To analyze whether there was a difference on the number of tasks accomplished according to participants' media and device usage, a series of Kruskal-Wallis tests were conducted considering the null hypothesis that there was no difference in the number of tasks accomplished by participants with different media and device usage.

Results indicate that according to the frequency participants' watch television, there is a difference (p value = 0.03) on the number of tasks on average they accomplished. Comparing the frequency participants watch television, according to parents' responses on the questionnaire, with the number of tasks participants accomplished it can be noticed that participants who watch TV more frequently accomplished more tasks than participants who watch TV occasionally (Table 10).

Frequency participants watch TV	Number of participants	Average (median) number of tasks accomplished
Several times per day	10	4.5
About once a day	9	3.0
A couple of times a week	3	1.0

Table 10. Average tasks accomplished by frequency participants watch TV

According to the computer usage, results do not indicate a statistically significant difference (p value = 0.269) on the average number of tasks accomplished among participants who use the computer more regularly or occasionally. Nevertheless, despite the fact that there is no statistically significant difference among participants indicating that computer usage have an effect on the number of tasks they accomplished, it can be noticed, from the Table 11 below, that participants who use the computer on a daily basis accomplished on average five out of the six tasks they were asked to accomplish, notably more tasks than participants who use the computer occasionally accomplished on average.

Frequency participants use the computer	Number of participants	Average (median) number of tasks accomplished
Several times per day	1	5.0
About once a day	5	5.0
A couple of times a week	5	2.0
About once a week	5	5.0
About a couple of times a month	2	3.5
About once a month	1	1.0
Less often	2	4.0
Never	1	3.0

Table 11. Average tasks accomplished by frequency of computer use

Parents were asked on the questionnaire if their children use the remote control while watching television. Their responses compared with the number of tasks on average accomplished by participants indicate that there is no significant difference (p value = 0.272) on the number of tasks accomplished by participants who use the remote, those who do not use it and participants who only use the remote with supervision. It can be noticed, however, from the Table 12 below, that those participants who use the remote control accomplished on average more tasks than those who do not use it or use the remote only with supervision.

Remote control usage	Number of participants	Average (median) number of tasks accomplished
Yes	13	4.0
Only with supervision	2	1.5
No	7	3.0

Table 12. Average tasks accomplished by participants' remote usage

Isolating only the tasks accomplished with the remote control, another Kruskal-Wallis test showed that there is no significant difference in the number of tasks accomplished for participants who use the remote control, those who use it only with supervision and participants who do not use the remote (p-value=0.189). The number of tasks accomplished using the remote by the thirteen participants whose parents stated they do use the remote varied between one and three (mean = 2). Only two participants' parents declared they use the remote control only with supervision, one of those accomplished one out of the three tasks while the other did not accomplish any task. Among the seven participants whose parents' response on the questionnaire was that they do not use the remote control, the number of tasks accomplished using the remote was between one and three, as the participants who use the remote only with a slightly lower mean (mean = 1.71). There is not enough evidence on video data either that elucidates a significant difference in the tasks accomplished by participants who have more experience with the remote control compared with those with less or no experience using the device.

A Kruskal-Wallis test was conducted to verify if the average number of tasks accomplished is different among participants' according to their mouse usage. The result indicates that there is no statistically significant difference (p value = 0.297) on the number of tasks accomplish for participants who use the mouse, those who do not use it and participants who only use the mouse with supervision. From the Table 13 below, however, it can be noticed that participants who use the mouse accomplished on average more tasks than participants who do not use the mouse. There is not enough data to provide evidence that participants who use the mouse with supervision accomplish more tasks because in the sample there is only one participant whose parents mentioned uses the mouse with supervision.

Mouse usage	Number of participants	Average (median) number of tasks accomplished
Yes	19	4.0
Only with supervision	1	6.0
No	2	2.5

Table 13. Average tasks accomplished by participants' mouse usage

In order to measure if there was a difference on the number of tasks accomplished with the mouse according to participants' mouse usage a Kruskal-Wallis test was conducted and indicates no evidence that participant's mouse usage affected the number of tasks accomplished with the mouse (p value=0.193). The small sample size could, however, have interfered with this result, since there were only two participants who do not use the mouse and one who uses it only with supervision. For this reason, it was decided to look at the video data for more detail. There were nineteen participants whose parents affirmed they use the mouse, but in this section the focus will be on the three other participants to investigate if their lack of experience with the device affected their interaction with the prototype application. The participant in Brazil whose parents said used the mouse only with supervision (Participant 2) accomplished the three tasks he was asked to using the mouse and the way he interacted with the prototype while accomplishing the tasks was very similar to that of participants whose parents affirmed use the mouse. The participant in Brazil (Participant 5) whose parents declared that she does not use the mouse accomplished only one out of the three tasks using the mouse as input device. She seemed to interact, however, as participants whose parents affirm they use the mouse and failed to accomplish two tasks because she clicked on screenshots others than the ones of the videos she was asked to find, but she managed to select the right category for the simple search and the two categories for the Boolean search using the mouse. The participant in the UK (Participant 12) whose parents stated she does not use the mouse did not accomplish any task using the mouse as input device. She struggled using the mouse, tried to use the touch pad but it did not seem to make the task any easier for her, so she was told to interact using the keyboard as an input device and accomplished one out of the three tasks using the keyboard as a substitute for the mouse.

In order to measure if participants' media and device usage had an effect on the time they took to accomplish tasks a series of Kruskal-Wallis tests were conducted with the null hypothesis that there was no difference on the time taken to accomplish tasks among the participants with different patterns for media and device usage.

Results indicate that according to the frequency participants' watch television there is no significant difference (p value = 0.441) on the time taken to accomplish tasks by participants who watch TV on a daily basis and those who watch it only a couple of times a week. The average time taken to accomplish tasks is similar for participants whose parents mentioned watch television several times per day, about once a day and a couple of times a week.

Regarding the frequency of computer use, another Kruskal-Wallis test indicates that there is no significant difference either on the time taken by participants to accomplish tasks according to the frequency they use the computer (p value = 0.121). The average time taken by participants who use the computer frequently is similar to the average time taken by participants who use the computer occasionally to accomplish tasks.

Results from a Kruskal-Wallis test suggest that there is no statistically significant difference on the time to accomplish tasks and participants' remote control usage (p value = 0.289). Participants whose parents said use the remote when watching television and participants whose parents mentioned do not use the remote or use it only with supervision, took on average the same time to accomplish tasks.

Nevertheless, if only the tasks accomplished using the remote are considered, then there is no statistical difference on the time taken for participants who use the remote, do not use it or only use it with supervision (p value = 0.630), but it may be noticed that the average (both mean and median) time taken to accomplish tasks using the remote control by participants who use the remote without supervision is slightly lower than the average time taken by participants who do not use the remote. Participants who use the remote accomplished tasks faster, they took on average (mean) 125.5 seconds to accomplish tasks with the remote while the participants who do not use the remote took on average (mean) 170.4 seconds to accomplish the tasks. It was not included on this last calculation the data from the two participants whose parents mentioned use the remote only with supervision, because one participant did not accomplish any task so the time to accomplish tasks could not be measured, while the other participant accomplished only one task what would not provide enough evidence of the effect using the remote with supervision could have on the time to accomplish tasks.

A Kruskal-Wallis test was carried out and indicates that the time to accomplish tasks is not statistically significant different for participants who use the mouse, those who use the device only with supervision and participants who do not use the mouse (p value = 0.801).

Considering only the tasks accomplished using the mouse, there is no significant difference either on the average time taken for participants who use the mouse, those who do not use it and participants who use the device only with supervision to accomplish tasks (p value = 0.678). The time taken to accomplish tasks by participants whose parents stated use the mouse, those who do not use and participants who only use it with supervision was similar.

A Kruskal-Wallis test indicates that there is a statistically significant difference on the number of hints given to participants according to the frequency they watch television (p value = 0.016). The test results suggest that participants who watch television less often were given on average, mean and median, fewer hints than participants who watch television more often. It has to be noted, however, that only three participants' parents affirmed watch television a couple of times a week, and those participants accomplished on average only one task.

Another Kruskal-Wallis test suggest that there is no statistically significant difference on the number of hints given and the frequency participants use the computer (p value = 0.207). The average number of hints given to participants who use the computer often was similar to the average number of hints given to participant who use the computer occasionally.

A Kruskal-Wallis test was carried out and indicates that there is no statistically significant difference on the number of remote navigation hints given to participants whose parents affirmed use the remote control and those who do not use it or use the device only with supervision (p value = 0.101). It may be noticed, however, from the boxplot below (Figure 48) that on average, both the mean and the median, fewer remote navigation hints were given to participants who use the remote than to participants who do not use the remote or use it only with supervision.

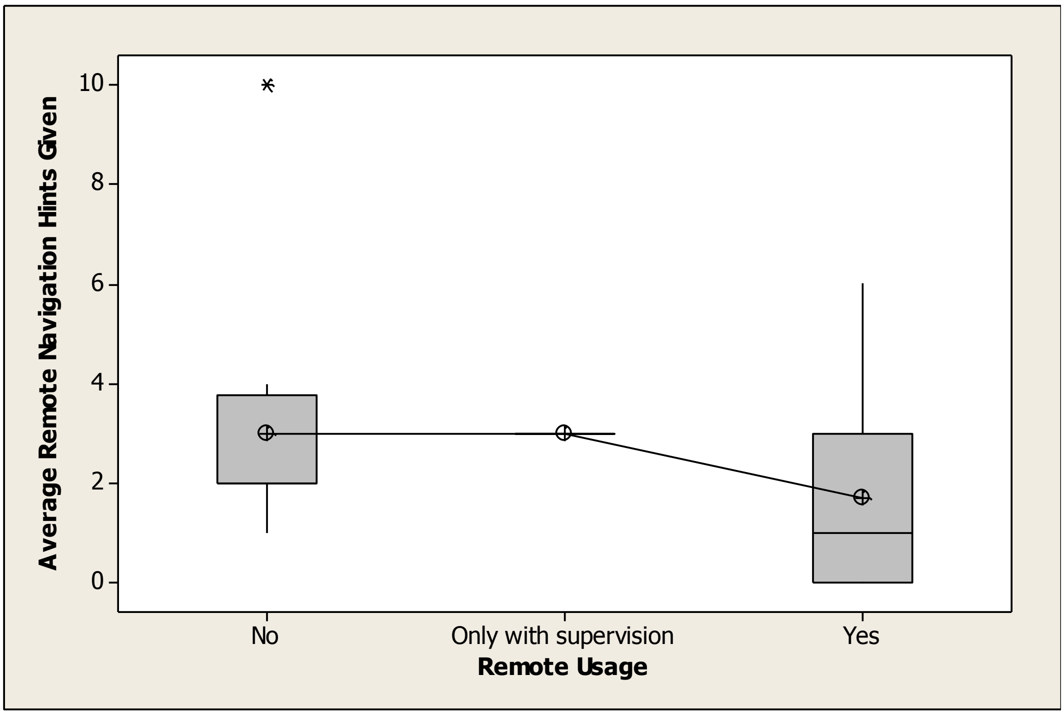


Figure 48. Average remote navigation hints by participants' remote usage

There is no evidence of statistically significant difference on the number of mouse navigation hints given to participants who use the mouse compared to those who do not use the device or use it only with supervision, a Kruskal-Wallis test indicates (p value = 0.253). It may, however, be noticed that on average, both mean and median, participants who use the mouse were given fewer mouse navigation hints than participants who do not use it or use it only with supervision (Figure 49).

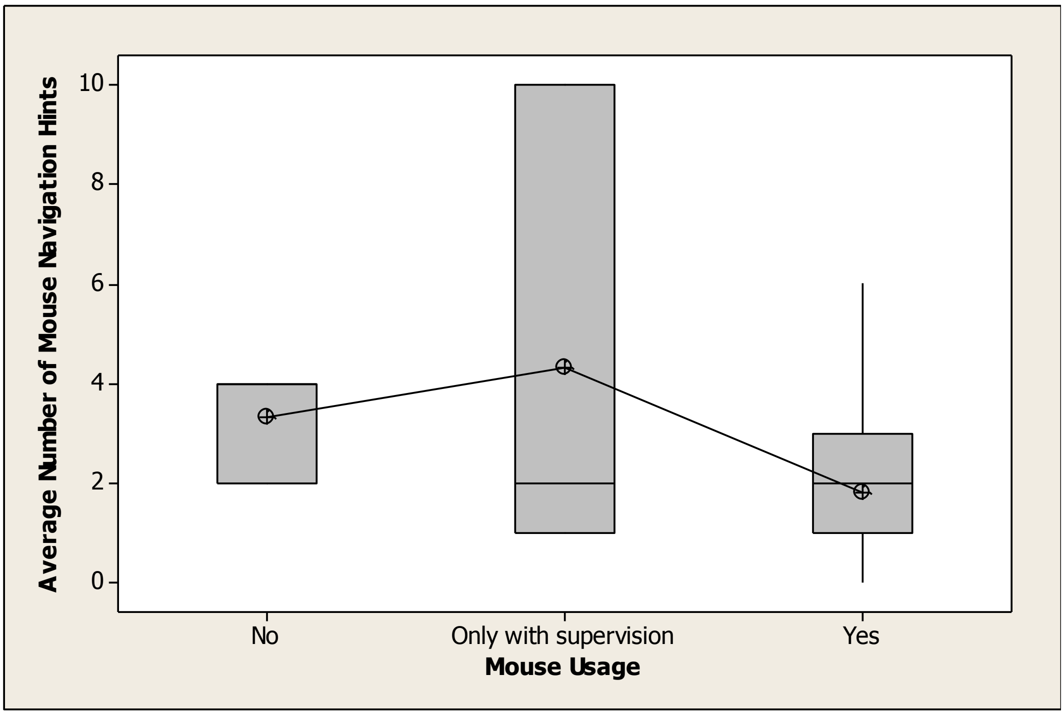


Figure 49. Average mouse navigation hints by participants' mouse usage

To investigate if participants' media and device usage had an impact on the number of interactions they performed to accomplish tasks a series of Kruskal-Wallis tests were carried out, in this case the null hypothesis was that there was no difference on the number of interactions performed to accomplish tasks among the participants with different patterns for media and device usage.

A Kruskal-Wallis test indicates that there is no statistically significant difference on the number of interactions performed to accomplish tasks by participants who watch television several times per day, about once a day and a couple of times a week (p value = 0.057).

Two other Kruskal-Wallis tests suggest no statistically significant difference on the median number of interactions performed by participants who use the computer frequently and those who use it occasionally (p value = 0.127).

Considering participants' remote usage, a Kruskal-Wallis test indicates that there is no statistically significant difference in the number of interactions performed to accomplish tasks by participants who use the remote, those who do not use it and participants who only use the remote with supervision (p value = 0.121). Regarding the number of buttons pressed only, there is no evidence of a statistically significant difference on the number of interactions performed by participants who use the remote unsupervised, those who do not use the device and participants who only use the remote with supervision (p value = 0.121).

A Kruskal-Wallis test suggests that there is a difference on the number of interactions performed to accomplish tasks by participants who use the mouse, those who do not use the device and participants who only use the mouse with supervision (p value = 0.024). Participants whose parents' affirm use the mouse unsupervised performed significantly fewer interactions to accomplish tasks than participants who do not use the mouse or use it only with supervision (Table 14).

Mouse usage	Number of participants	Average (median) number of interactions performed
Yes	19	40
Only with supervision	1	135
No	2	192

Table 14. Average interactions performed by participants' mouse usage

Considering only the clicks performed to accomplish tasks, instead of the total number of interactions, a Kruskal-Wallis test indicates that there is no statistically significant difference on the number of clicks performed by participants who use the mouse, those who do not use it and participants who only use the device with supervision (p value = 0.389). It may, however, be noticed that the average number of clicks performed to accomplish tasks by participants who use the mouse is lower than the average number of clicks performed by the participant who use the mouse supervised, that, in turn, is significantly lower than the average number of clicks performed by the two that do not use the mouse (Table 15).

Mouse usage	Number of participants	Average (median) number of clicks performed
Yes	19	7
Only with supervision	1	18
No	2	89.5

Table 15. Average clicks performed by participants' mouse usage

5.3.2.6 Input Device

To analyse the effect the input device could have had on the number of tasks accomplished, a Wilcoxon[4] signed rank test was carried out and indicated (p value=0.590) that the null hypothesis should not be rejected, there is no difference in population median of tasks accomplished with the mouse and the remote control, on average a similar amount of tasks was accomplished using the mouse and the remote.

Among the three tasks participants were asked to accomplish with each device, from Figure 50 below it can be noticed that the average, both mean and median, of the number of tasks accomplished with the mouse and the number of tasks accomplished

[4] The Wilcoxon signed rank test is a non-parametric test for repeated measurements on a single sample used to calculate the median of a population of differences (Rees, 2001).

with the remote control are very similar. There is, however, slightly more variation in the number of tasks participants accomplished with the mouse than with the remote.

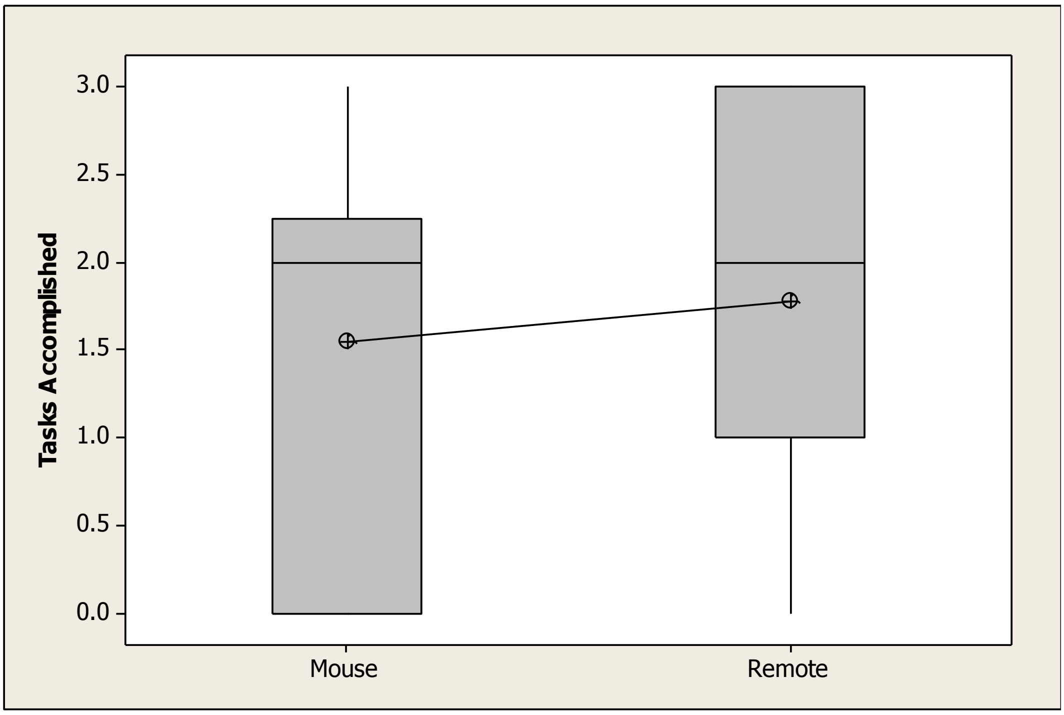


Figure 50. Average tasks accomplished by input device

The two additional input devices that were alternative to the mouse, the touch pad and keyboard, were not included in the graph or calculations because they were only used by a very small number of participants. After watching the tutorial with instructions for the interaction with the prototype, a participant in the UK (Participant 19) was asked to accomplish the first task using the mouse, she pointed to the computer touch pad, however, and said *my mommy and daddy got that one*. She was then told she could interact using the mouse or the touch pad and chose to use the touch pad. Two other participants in the UK (Participants 12 and 16) were offered an alternative input device because they were struggling to interact using the mouse, both were told to use the touch pad. They tried to use it, but it did not make the interaction any easier so they were told they could use the keyboard instead. In total three participants used an alternative input device instead of the mouse, the participant who used the touch pad accomplished the three tasks and the participants who used the keyboard accomplished one out of the three tasks with that device. Since one participant (Participant 19) chose not to accomplish any tasks with the mouse and two participants (Participant 12 and 16) were not able to accomplish any task using the mouse, in the calculations and graphs involving input devices data, for these three participants who used an alternative input device the number of tasks accomplished with the mouse were equivalent to zero.

Participants accomplished tasks on average slightly faster using the mouse than using the remote (Figure 51). Alternative input devices were discarded from these analyses as well because they were used by only three participants. The participant who used the touch pad accomplished the three tasks and took on average 187 seconds to accomplish each task, which is above the average participants who used the mouse took. The two participants who used the keyboard accomplished only one task and took 557 and 315 seconds respectively to accomplish the task, average time well above the average participants who used the mouse took.

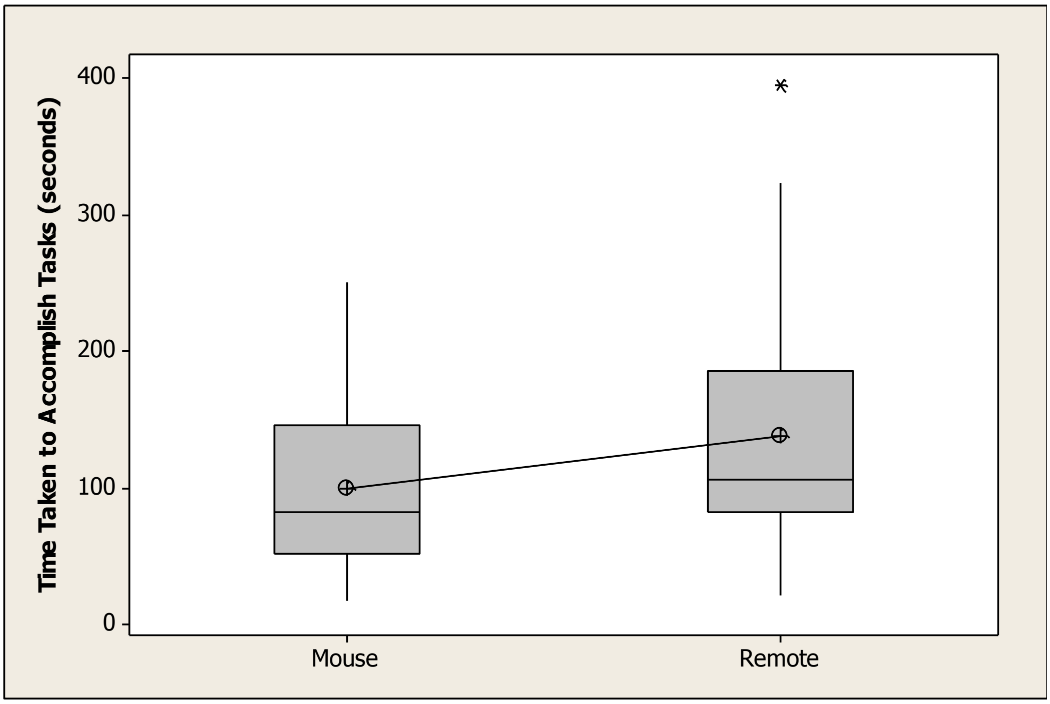


Figure 51. Average time taken to accomplish tasks by input device

On average, a larger number of hints were given to participants to accomplish tasks using the remote (mean = 2.16) than to accomplish tasks with the mouse (mean = 1.18) and there is more variation on the number of hints related to remote navigation than on those concerning the mouse (Figure 52).

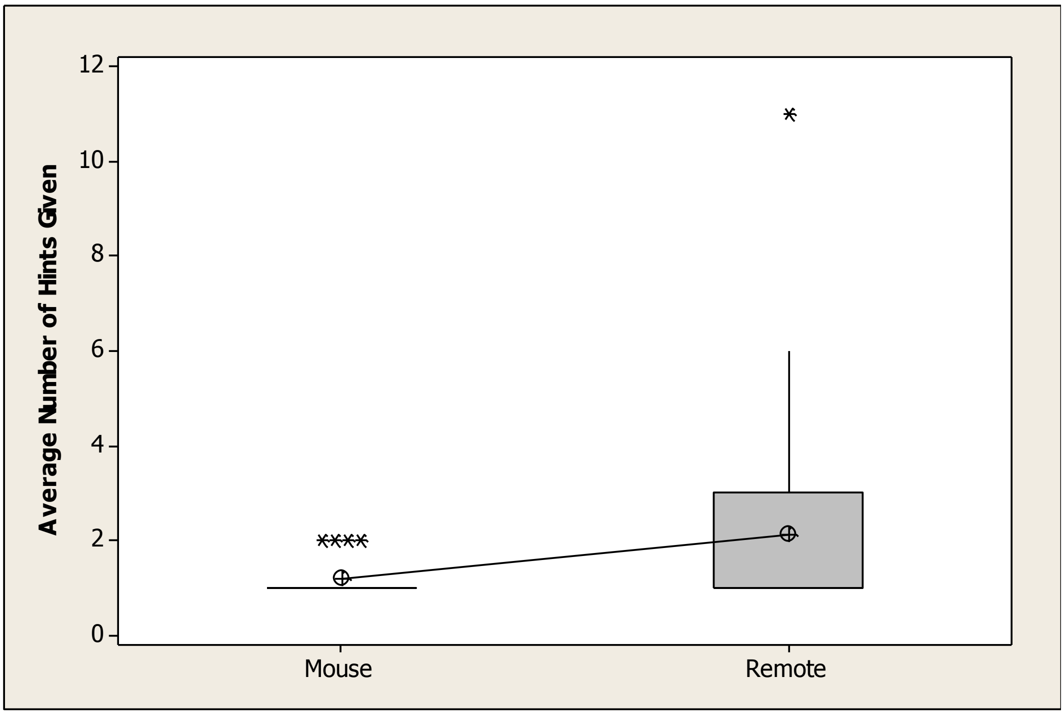


Figure 52. Average hints given by input devices

For the three participants who used an alternative input device instead of the mouse hints for interaction were also given. The two participants who used the keyboard were given on average, both mean and median, two keyboard navigation hints (H15 and H16) to accomplish tasks using the device. The participant who used the touch pad was given on average one touch pad navigation hint (H13 and 14) to accomplish tasks.

The number of clicks performed with the mouse and buttons pressed on the remote control were calculated and compared to evaluate the number of interactions needed to accomplish tasks using each device.

The average, mean and median, number of interactions performed to accomplish tasks using the mouse is significantly lower and presents significantly less variation than the average number of interactions performed to accomplish tasks using the remote (Figure 53).

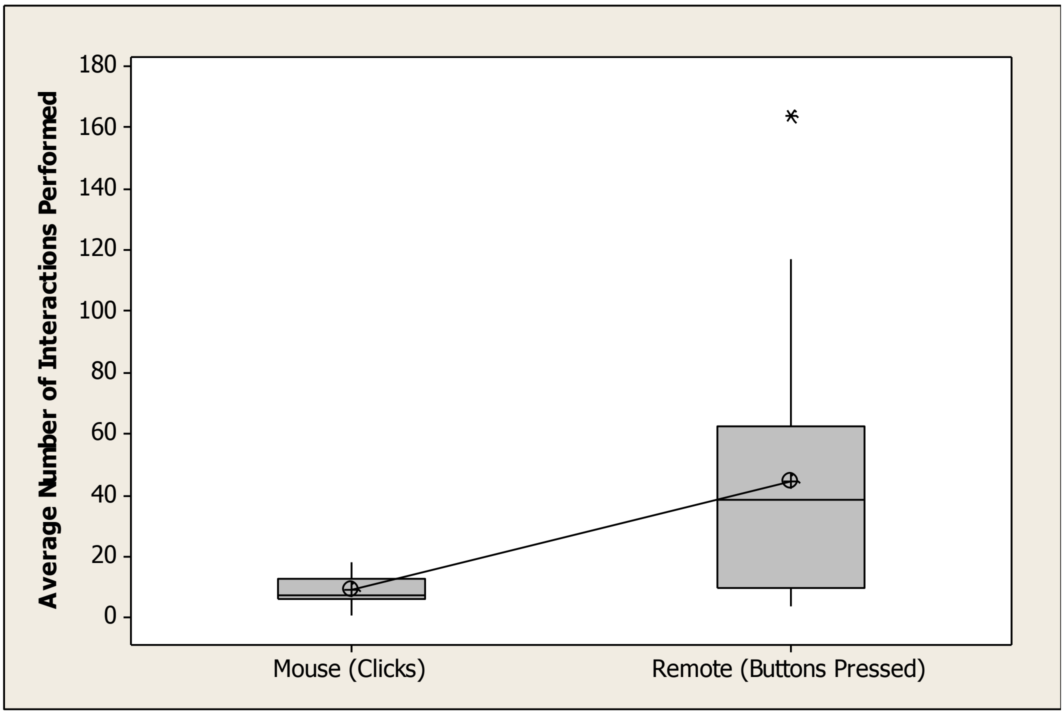


Figure 53. Average interactions performed by input device

The interactions performed by the three participants who used alternative input devices were not included on the graphs and calculations above. Participants 12 and 16, who accomplished one task using the keyboard as alternative for the mouse, performed 177 and 122 interactions respectively, significantly more than participants who interacted using the mouse. Participant 19 who accomplished three tasks using the touch pad as alternative for the mouse needed on average 3.66 interactions to accomplish each task, less than the average number of interactions participants who used the mouse needed.

As reported previously, participants' age was found to have an effect on their interaction with the prototype, older participants accomplished more tasks, thus for subsequent analysis participants were divided into two groups, three year olds and four year olds. There were in total ten participants three years old and twelve participants four years old. Considering participants' age, it can be seen from the graph below (Figure 54) that the average of number of tasks accomplished with mouse and remote is similar. As discussed previously, older participants accomplish more tasks, four year olds accomplish more tasks on average than three year olds, but within each age, the average number of tasks accomplished with both devices is similar. Participants who are three years old accomplish on average (mean) 1.1 tasks with the mouse and 1.4 with the remote, while four years old participants accomplish on average (mean) 1.9 tasks with the mouse and 2.1 with the remote control. It can be noticed that, especially for three years old participants, there is slightly more variation on the average number of tasks using the mouse than those using the remote. While only one three year old participant and one four year old participant did not accomplish any task with the remote, five three year olds did not accomplish any task using the mouse and one four years old participant, who chose to use the touch pad, did not accomplish any task with the mouse.

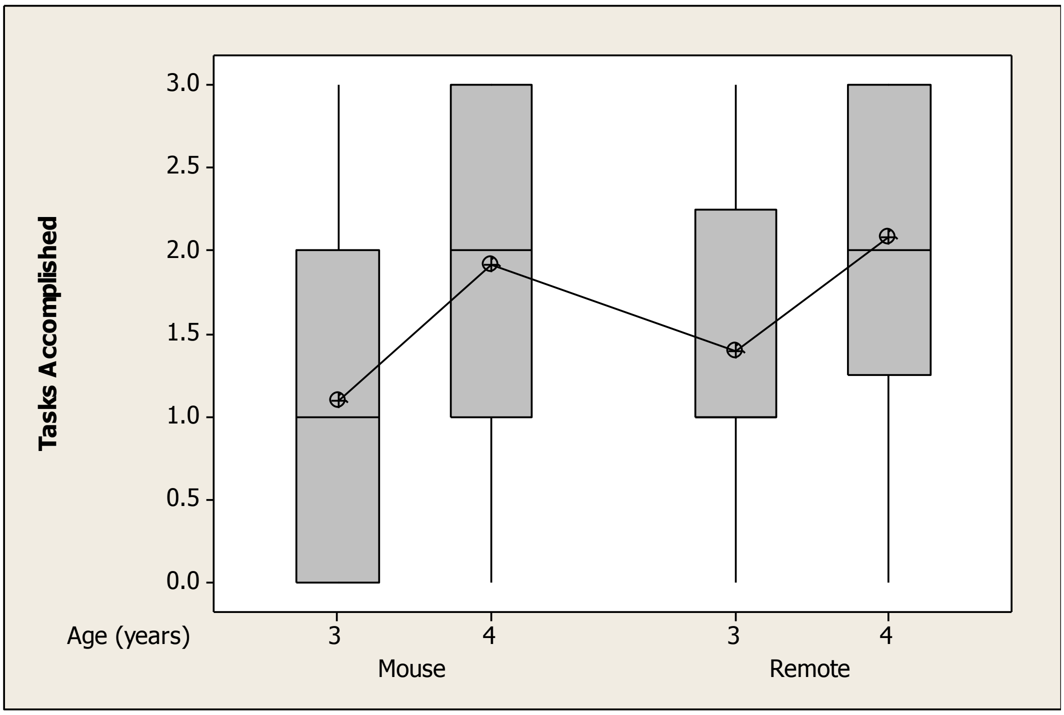


Figure 54. Average tasks accomplished by age and input device

There is not enough data to provide evidence that the time taken to accomplish tasks using each device is different according to participants' age.

Considering hints given to accomplish tasks, it may be noticed that on average (mean) a similar amount of hints relating both mouse and remote navigation (H1 – H8) were given for three and four years old participants (Figure 55). For participants who used the alternative input devices the results are similar. Both participants who used the keyboard were three years old and given on average two navigation hints, between the average of hints given for navigation using mouse and those for navigation with the remote for the same age group. The participant who chose to use the touch pad was four years old and was given on average one touch pad navigation hint, slightly less than the average number of hints given for mouse navigation to participants in the same age group.

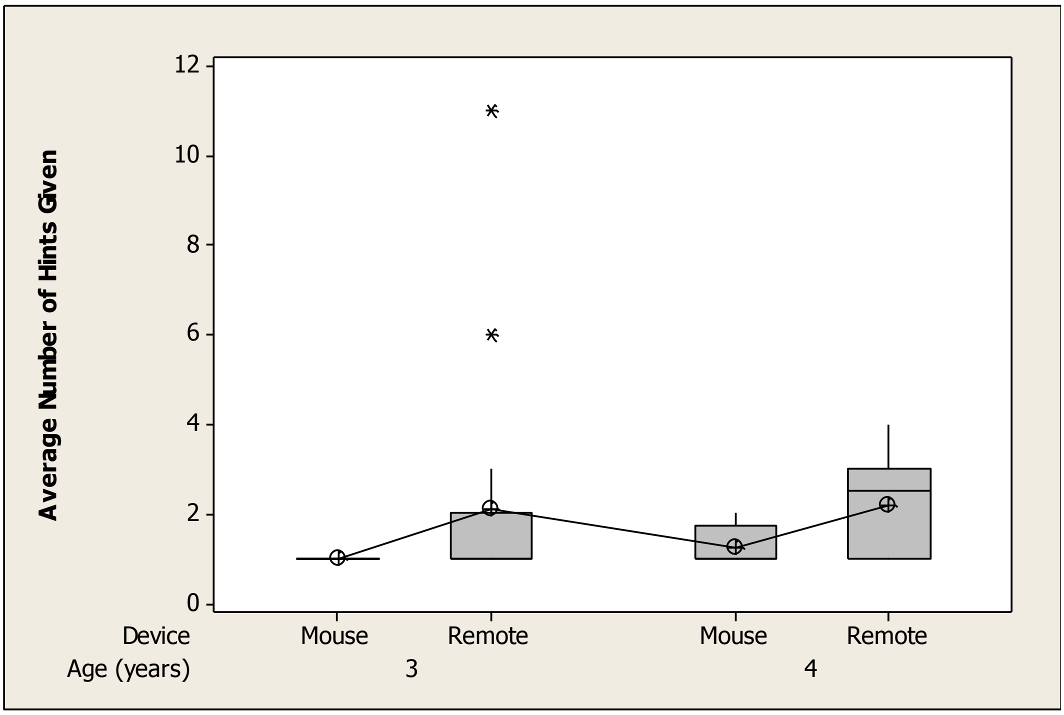


Figure 55. Average hints given by input device and age

The number of interactions three years old participants performed using the mouse, clicks, and the remote control, buttons pressed, appears to be very similar. Four years old participants, on the other hand, performed significantly fewer interactions using the mouse than using the remote in order to accomplish tasks (Figure 56).

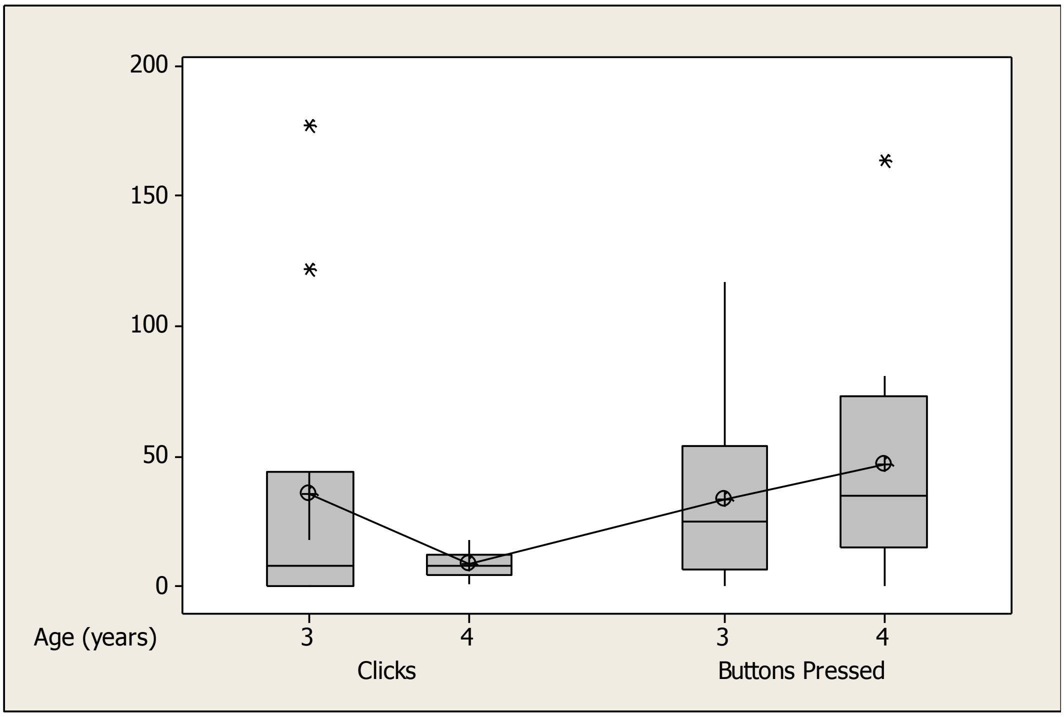


Figure 56. Number of interactions by participants' age

There is not enough evidence that the number of tasks accomplished using each device is significantly different for boys and girls.

Participants from both genders seem to accomplish tasks using the mouse faster than using the remote, but the time on average taken by boys and girls to accomplish tasks using both devices is very similar (Figure 57).

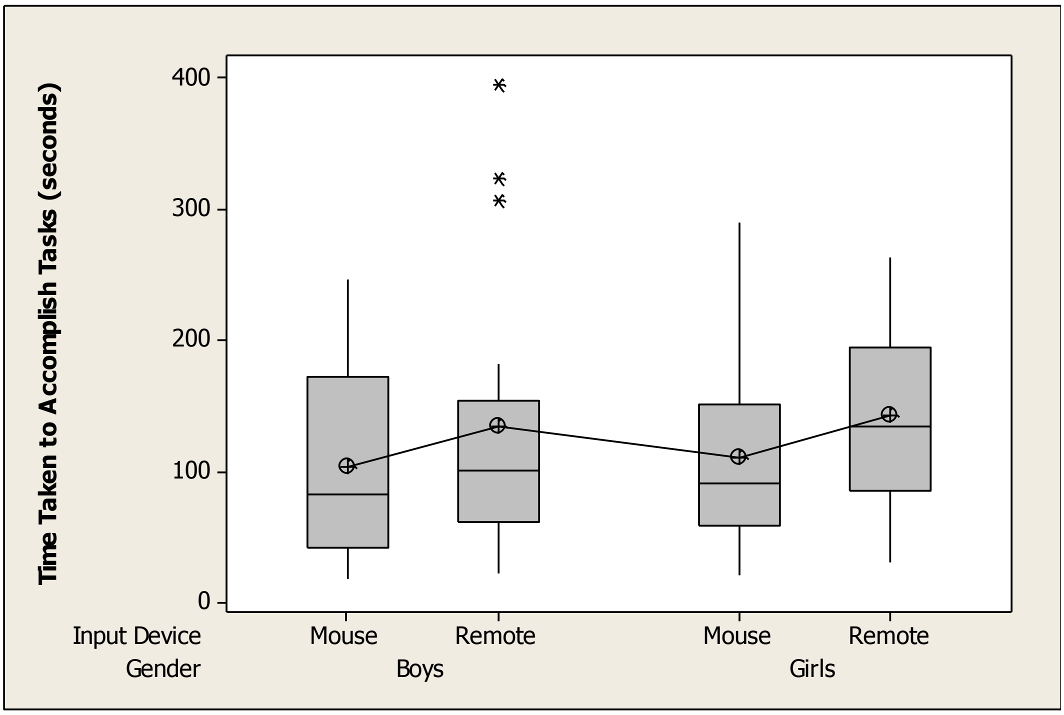


Figure 57. Average time taken to accomplish tasks by device and gender

There is not enough evidence that the number of hints given for girls and boys to assist on the navigation using each device is significantly different.

Participants of both genders performed significantly fewer interactions to accomplish tasks using the mouse than using the remote control. There is, however, no significant difference on the number of interactions performed with each device by boys and girls.

There is not enough evidence that the number of tasks accomplished using each device is significantly different for participants in the two country contexts. Participants in Brazil seemed to have accomplished on average similar amount of tasks using the mouse and the remote. Participants in the UK also seemed to have accomplished similar amount of tasks with each device. Comparing both country contexts, however, there is a difference in the number of tasks accomplish using each device. The number of tasks participants in Brazil accomplished using the remote control is significantly different from the number of tasks participants in the UK accomplished using the remote (p value = 0.0078). Participants in Brazil accomplished on average (mean) 2.4 tasks using the remote while participants in the UK accomplished on average 1.2. According to tasks accomplished using the mouse, there is also a significant different among the tasks accomplished by participants in Brazil and in the UK (p value = 0.0047). Participants in Brazil accomplished on average (mean) 2.3 tasks whereas participants in the UK accomplished 0.8 tasks.

There is not enough evidence that the time participants in Brazil and in the UK took to accomplish tasks using the mouse and the remote is statistically significant different (Figure 58). As discussed previously, participants took on average less time to accomplish tasks using the mouse than using the remote. However, there is no significant difference on the time participants in Brazil and in the UK took to accomplish tasks using the mouse or the remote.

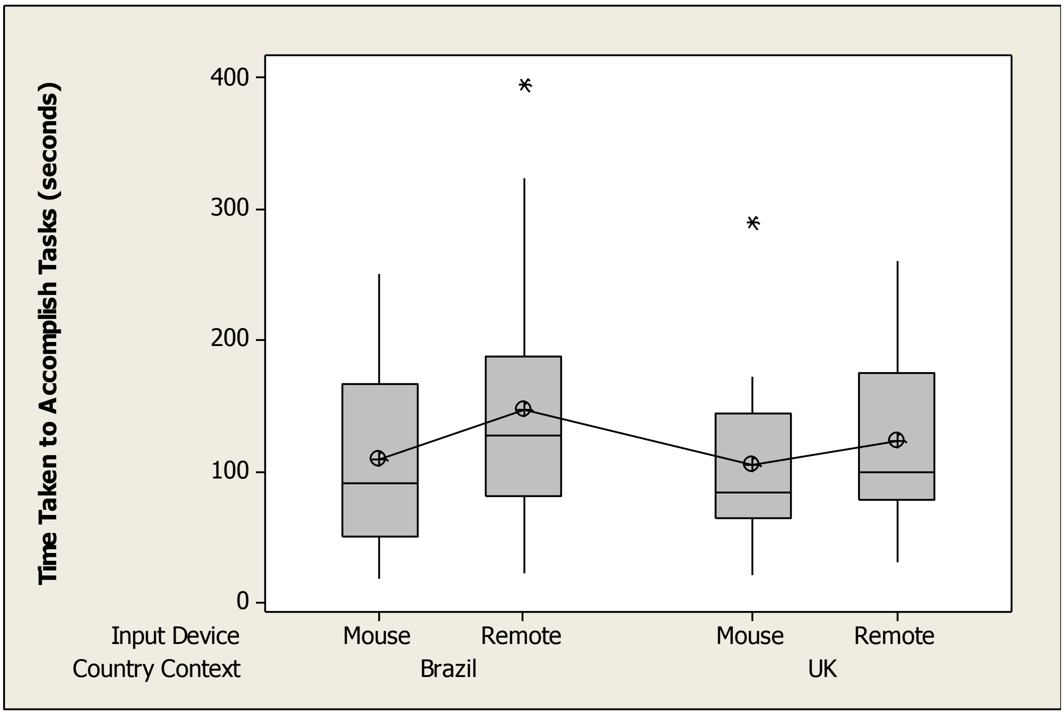


Figure 58. Time taken to accomplish tasks by input device and country context

There is not enough evidence that there is a statistically significant difference on the number of hints given for participants in Brazil and in the UK to assist on the navigation using the mouse or the remote control.

Participants in Brazil performed significantly more interactions to accomplish tasks using the remote than using the mouse. The difference on the number of interactions with each device was not as significant for participants in the UK (Figure 59). Comparing participants in both countries it may be noticed that the average number of interactions (clicks) to accomplish tasks with the mouse is slightly lower for participants in Brazil than for those in the UK. For tasks accomplished using the remote, however, participants in Brazil performed on average significantly more interactions (buttons pressed) than participants in the UK.

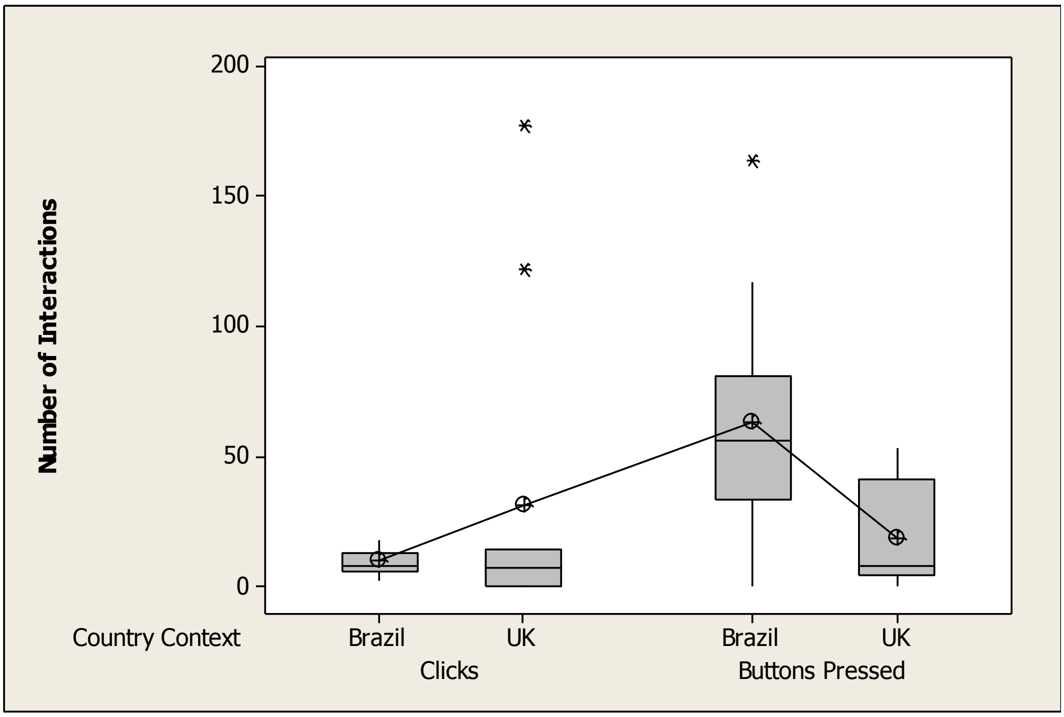


Figure 59. Number of interactions by country context

There is not enough evidence that there is a statistically significant difference on the average number of tasks accomplished by participants on each condition using the remote or the mouse.

It can be noticed, however, that there is a difference on the time taken to accomplish tasks with each device for the four conditions (Figure 60). Participants on conditions one and three who have used the mouse first and remote second took longer to accomplish tasks with the mouse than with the remote. Participants on conditions two and four, on the other hand, who have used the remote first and the mouse second took longer to accomplish tasks with the remote than with the mouse. The difference is not statistically significant due to the size of the sample, but it was found worth reporting.

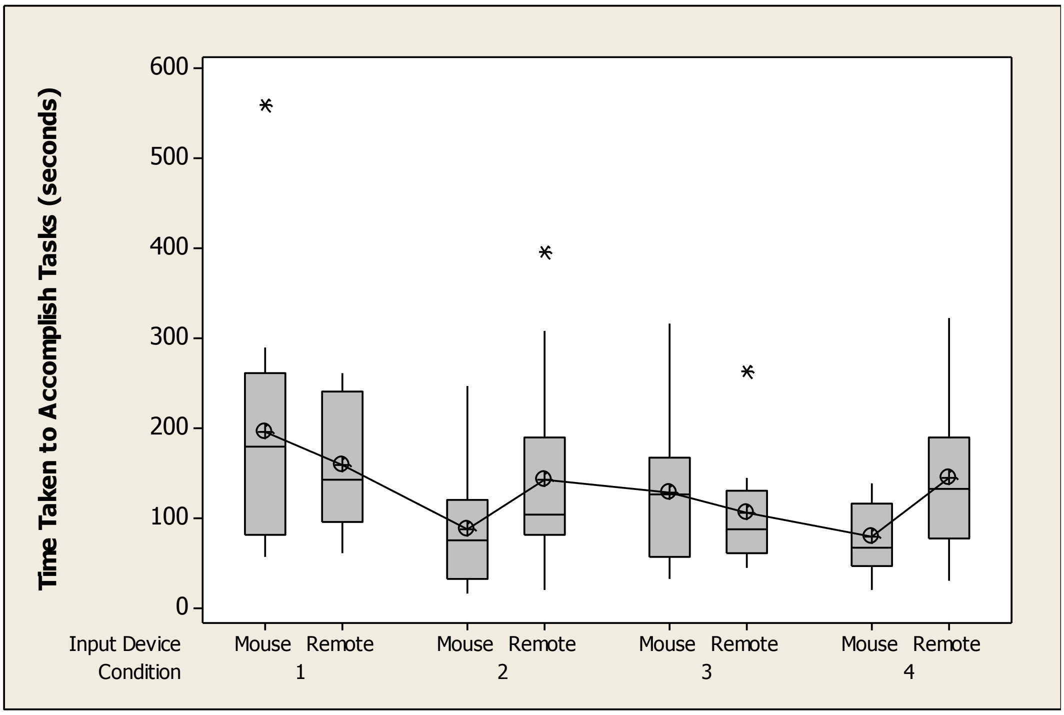


Figure 60. Average time to accomplish tasks by input device and condition

There is not enough evidence that a statistically significant different average number of hints were given for participants to navigate using the mouse or the remote in the four conditions.

There is not enough evidence either that there is a statistically significant difference on the number of interactions performed to accomplished tasks among the four different conditions for the mouse or the remote.

5.3.2.7 Task Type

A Friedman test[5] was carried out to verify if the task type influenced the number of tasks accomplished. The result (p value<0.001) indicates that the null hypothesis suggesting that the groups of data are equal should be discarded. There is a significant difference on the number of tasks accomplished for task types.

The number of tasks accomplished for task one, in which participants were asked to select a video already in display, appears to be significantly higher than the number of tasks accomplished for tasks two and three (Figure 61). However, the number of tasks accomplished for task two, in which participants were asked to perform a simple search selecting one category, appears to be similar to the number of tasks accomplished for task three, in which they had to select two categories.

[5] The Friedman test is a non-parametric test used to evaluate the differences between three or more treatment conditions in one sample, in order to test the null hypothesis: there is no difference between treatments, against the alternative hypothesis: at least one treatment has a different location from the others (Sprent, 1990).

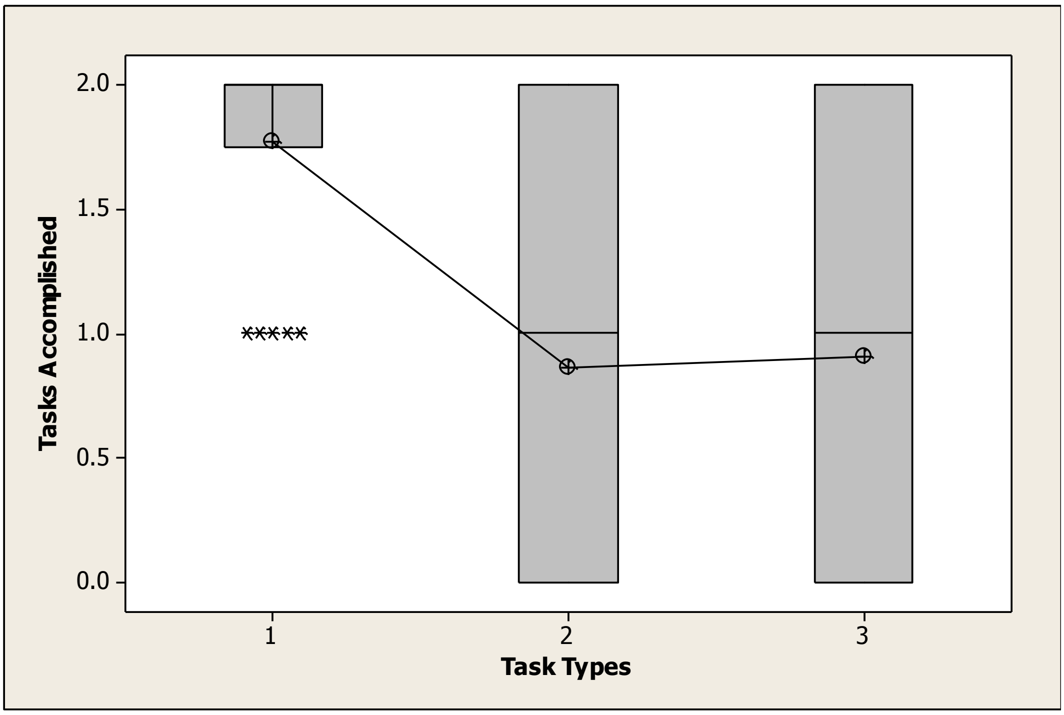


Figure 61. Accomplished tasks on average by task type

Regarding the time taken to accomplish tasks, it can be noticed that task one, in which participants were asked to select a video already in display, was accomplished significantly faster than tasks two and three, in which participants had to select one and two categories respectively to find the video they were asked to (Figure 62). The average time participants took to accomplish tasks two and three was similar, there was, however, more variation on average time to accomplish task two than to accomplish task three.

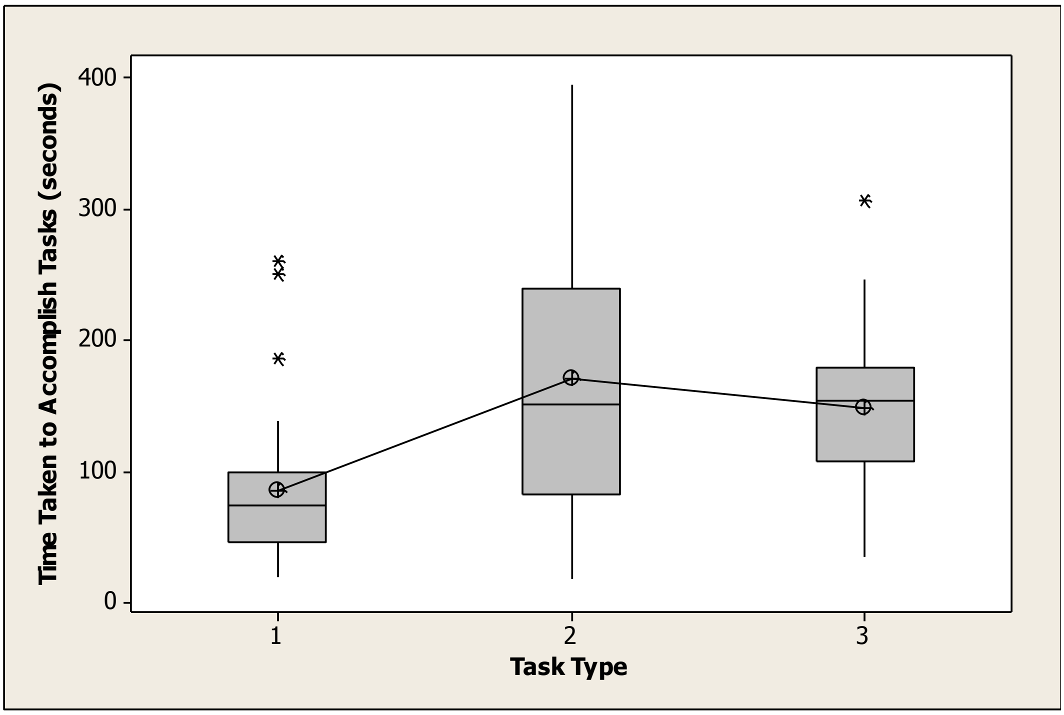


Figure 62. Average time to accomplish tasks by task type

According to the number of hints given, on average, both mean and median, a smaller number of hints were given to participants to accomplish task one compared with a larger number of hints given to participants to accomplish tasks two and three (Figure 63).

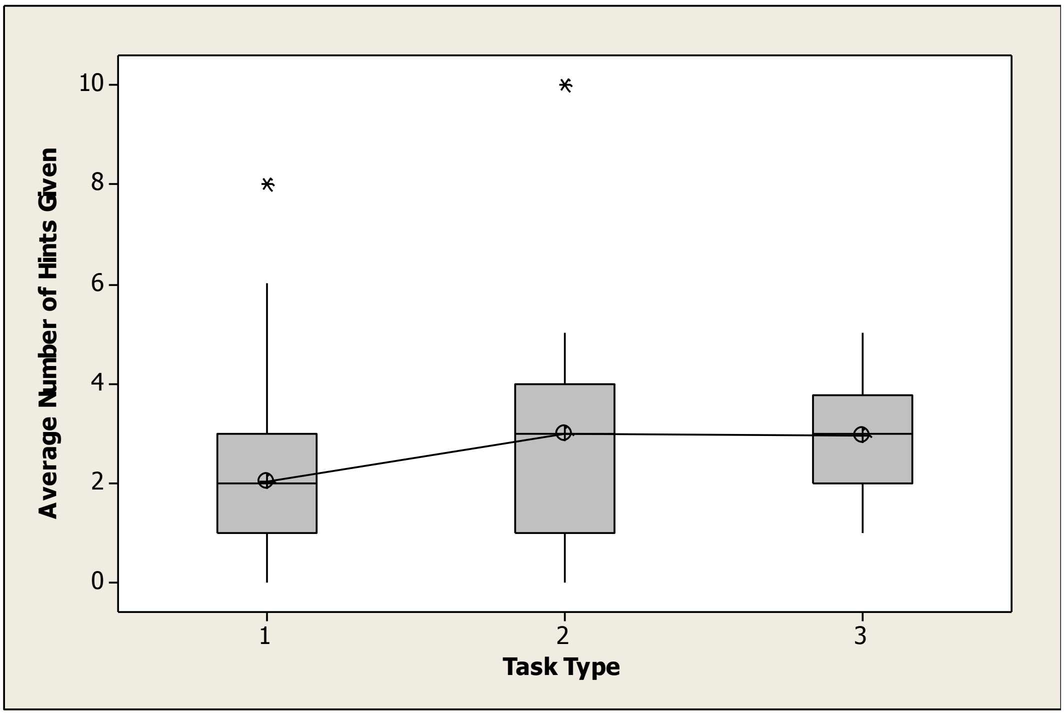


Figure 63. Average hints given by task type

It may be noticed that the average (median) number of interactions performed to accomplish tasks of the type one, in which participants had to select a video already in display, is lower and presents less variation than those for tasks two and three, in which participants had to select one and two categories respectively in order to accomplish the tasks (Figure 64).

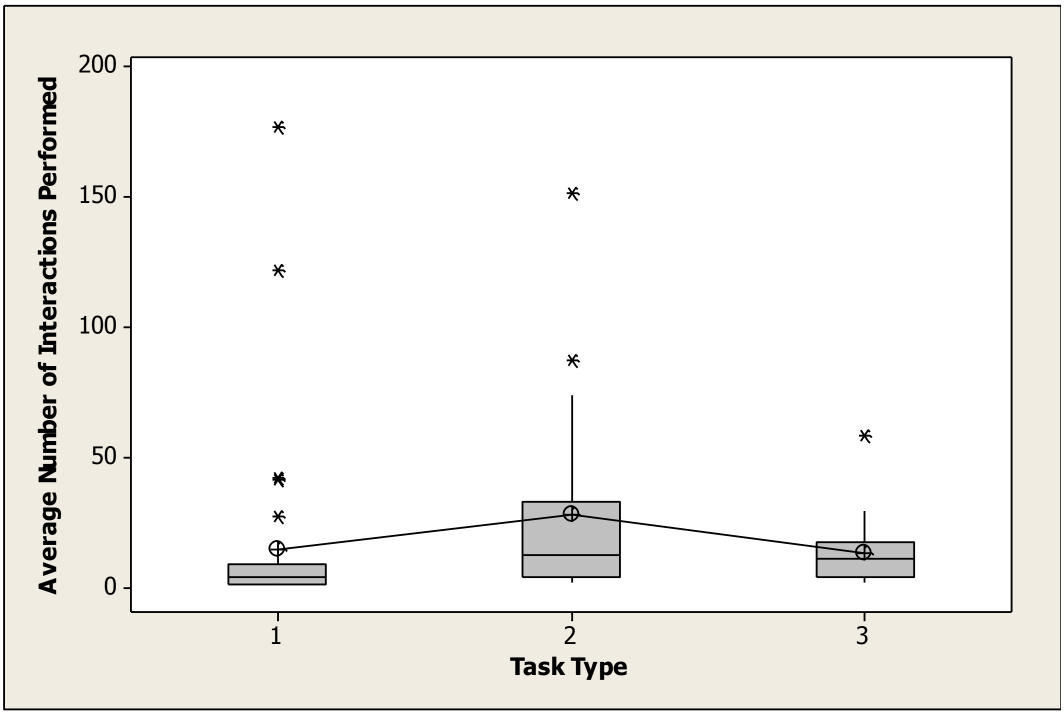


Figure 64. Average number of interactions by task type

According to participants' age and the type of the tasks accomplished, it can be noticed that for tasks one and two there is not a significant difference on the average number of tasks accomplished for the two age groups. Nevertheless, for task three there seem to be a difference on the number of tasks accomplished by three and four year olds, four years old participants accomplished on average more type three tasks than three years old participants (Figure 65).

For type one task, in which participants had to select a video on display, the average (mean) of tasks accomplished by three year olds was 1.6 while the average for the same task type for four year olds was 1.9. For task two, in which participants had to select one category in order to find the video, three year olds accomplished on average (mean) 0.6 tasks while four year olds accomplished on average 1.0 task. For task three, in which participants had to perform a Boolean search, the number of tasks on average (mean) accomplished by three year olds was 0.5 and four years old participants accomplished on average 1.25 tasks.

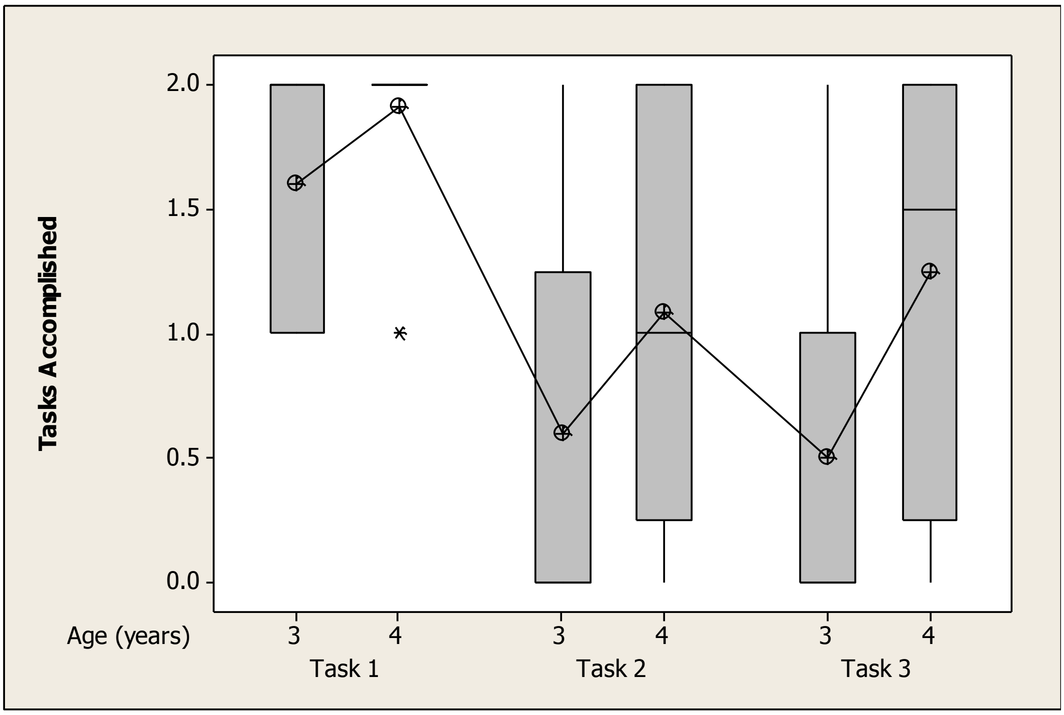


Figure 65. Average of tasks accomplished by task type and age

The average time participants in both age groups took to accomplish each task was also measured. For task one, in which participants had to select a video already being displayed in the 3D wheel, three year olds took more time to accomplish the task and there was more variation in the time they took compared with four year olds. For task two, however, in which participants were asked to perform a simple search selecting one category, the time taken to accomplish the task was similar for three and four year olds. For task three, in which participants were asked to perform a Boolean search, three year olds took more time and had more variation on the time taken to accomplish the task than four year olds. Those differences and similarities among the time taken to accomplish each task type according to participants' age may be observed on Figure 66 below.

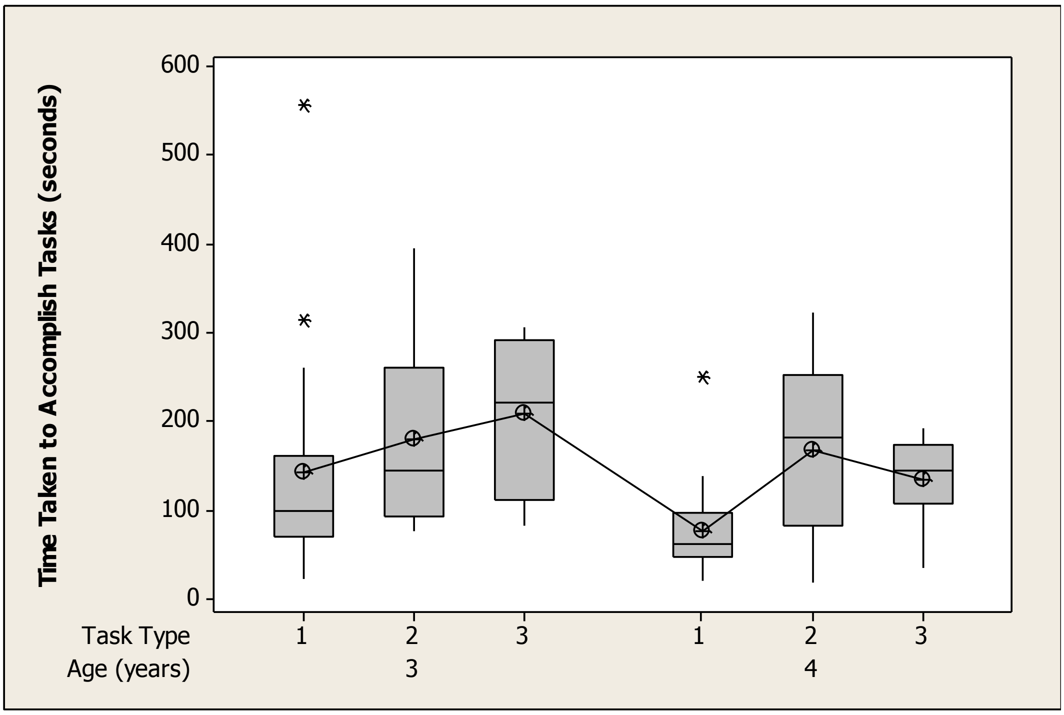


Figure 66. Time to accomplish tasks by task type and age

According to the boxplot below (Figure 67), it can be noticed that the average (mean) number of hints given to three years old participants is slightly higher than the number of hints given to four years old participants for tasks of the type one and two. Younger participants were given more hints than older participants. For tasks three, however, in which participants had to select two categories to find a video, the average number of hints given for participants of both ages was very similar.

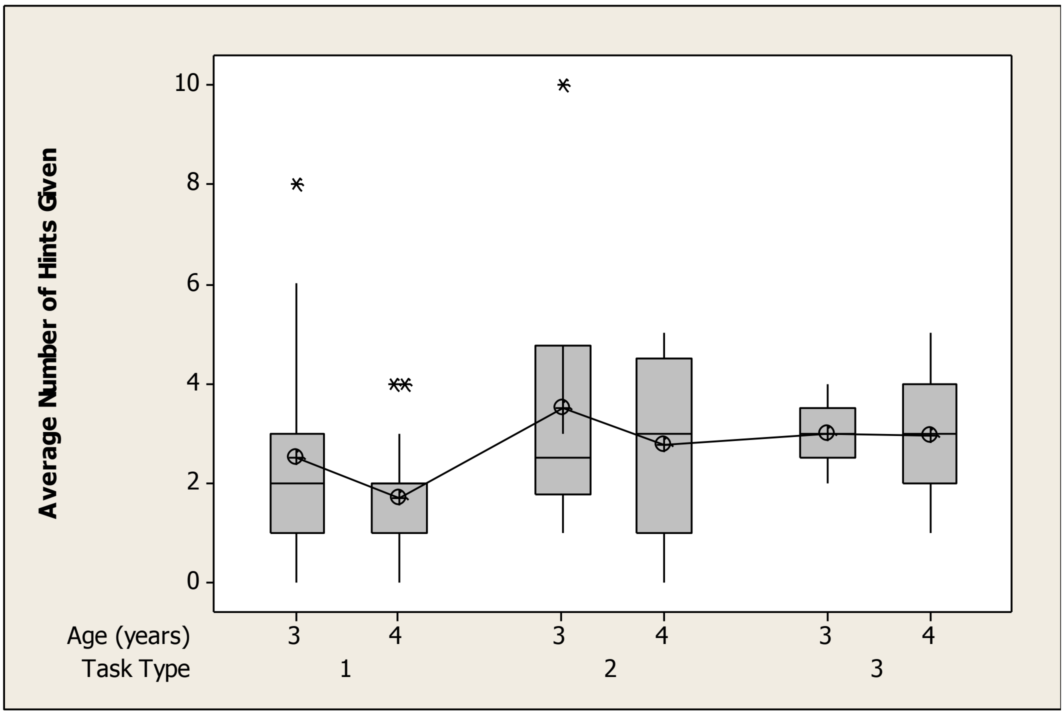


Figure 67. Average hints given by task type and age

It may be observed that on average (mean) a larger number of interactions were performed by three year olds to accomplish tasks one, in which they had to select a video already on display, compared with those performed by four year olds. For tasks two and three, in which participants had to select categories to find the video, the average (mean) number of interaction performed to accomplish the tasks by three and four years old participants is similar (Figure 68).

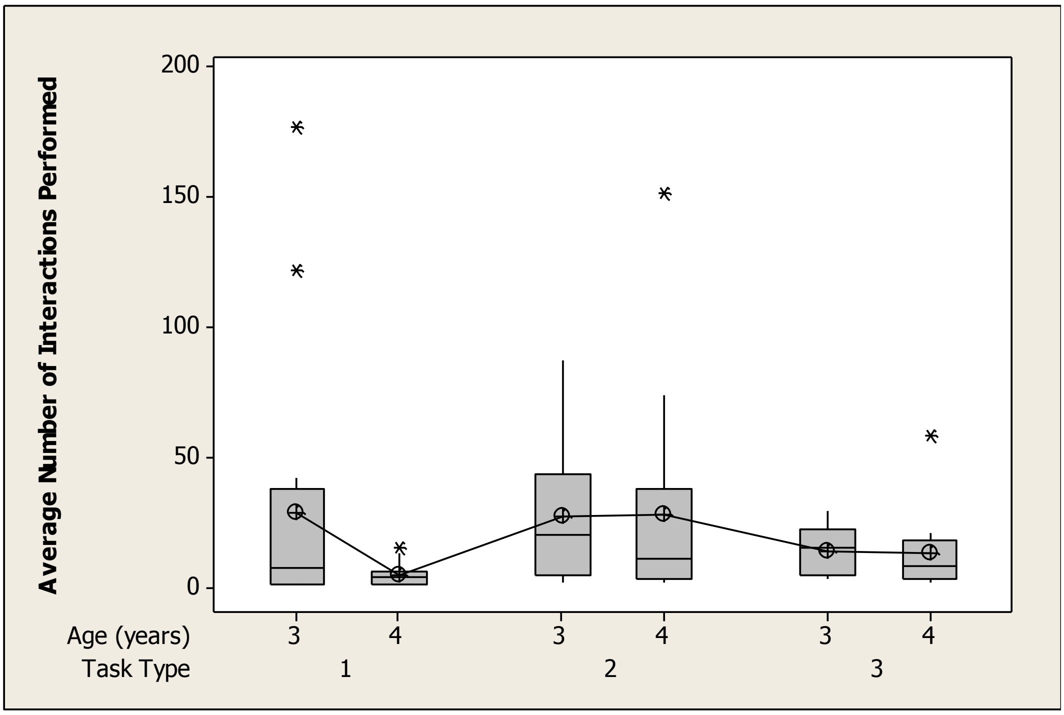


Figure 68. Average number of interactions by age and task type

There was not enough evidence that the number of tasks accomplished by each task type, time taken to accomplish each task, number of hints given, and number of interactions performed is significantly different among the two genders.

There is no evidence of a statistically significant difference on the number of type one tasks accomplished by participants in Brazil and in the UK. For task of the type two and three, however, there was a difference on the number of tasks accomplish among the two country contexts. As it may be seen from the Figure 69 below, for task one, participants of in both countries accomplished on average a similar number of tasks while for tasks two and three, participants in Brazil accomplished on average more tasks than participants in the UK.

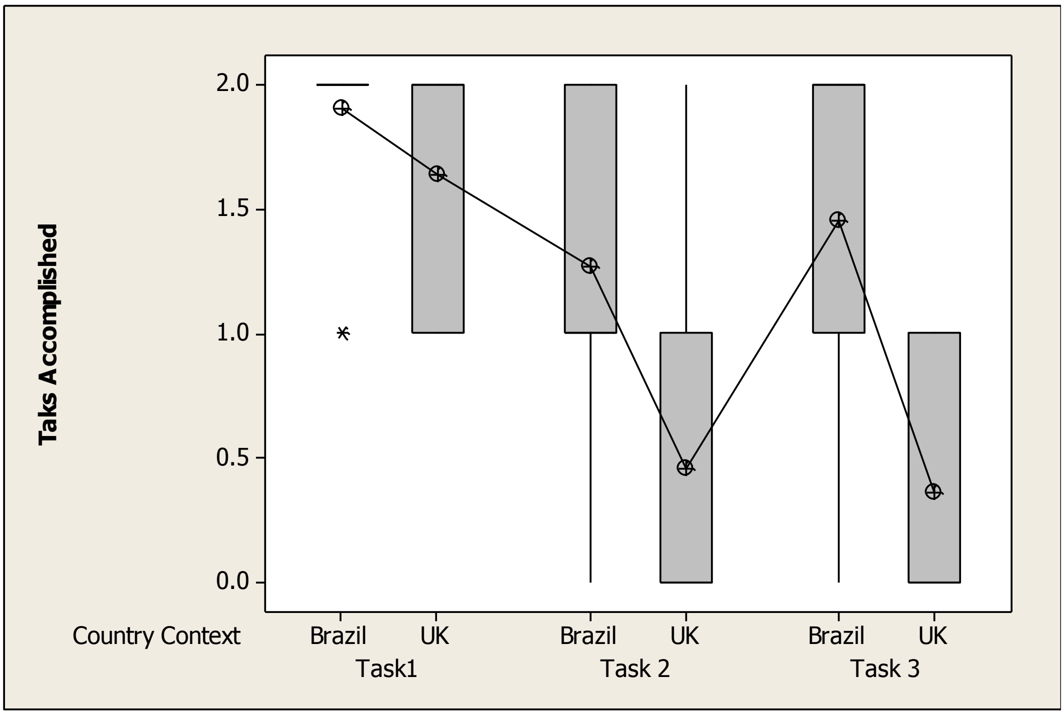


Figure 69. Accomplished tasks on average by task type and country context

There is not enough evidence that the time taken to accomplish tasks of each type was different among the two country contexts. The time taken by participants in Brazil and in the UK to accomplish tasks is similar for task one, two and task three. It is important to note, however, that the sample sizes in this case are different, especially for tasks two and three, since a significant larger number of participants in Brazil accomplished these tasks compared to participants in the UK (Figure 69).

There is not enough evidence either that there is a significant difference on the number of hints given to participants in Brazil and in the UK to accomplish tasks of the type one, two or type three. For tasks of the type one, in which participants had to select a video already being displayed, on average (mean) a slightly larger number of hints were given to participants in the UK compared to the number of hints given to participants in Brazil. For tasks of the type two, in which participants had to select a category to find a video, the average (mean) number of hints given to participants in both countries was similar. For tasks of the type three, in which participants had to select two categories, on average a slightly larger number of hints were given to participants in Brazil compared to those in the UK (Figure 70).

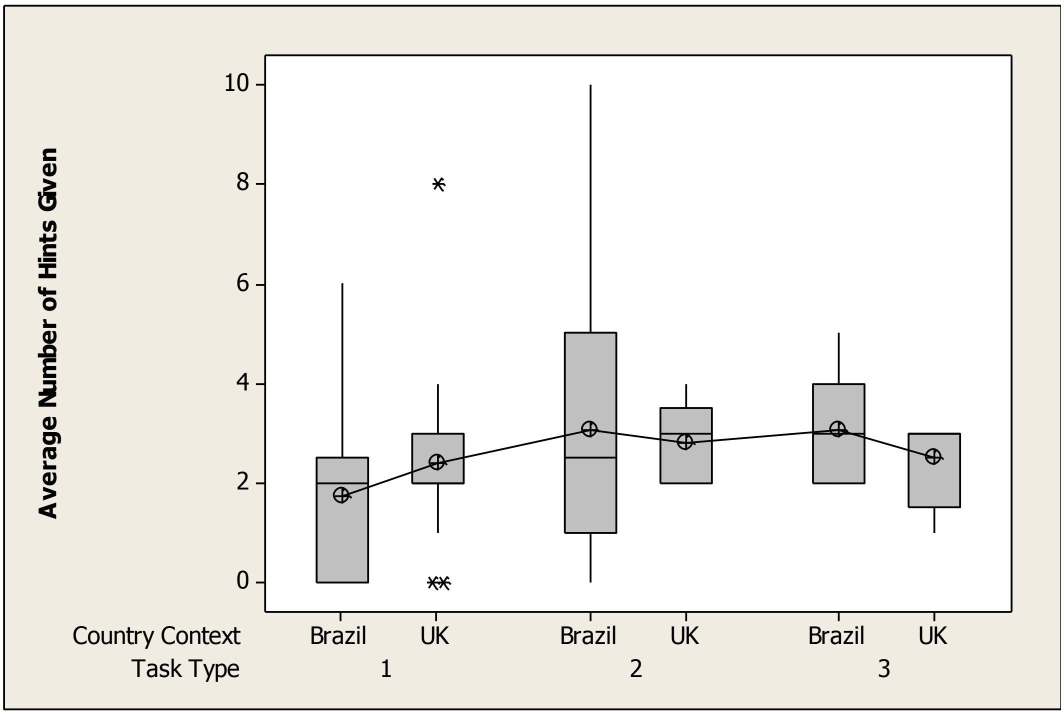


Figure 70. Average hints given by task type and country context

Considering the interactions, it may be seen from the boxplot below (Figure 71) that for task type one participants in the UK performed on average more interactions to accomplish the tasks than participants in Brazil. For tasks two and three participants in Brazil performed more interactions than participants in the UK, but the difference on average number, especially median, is not significant.

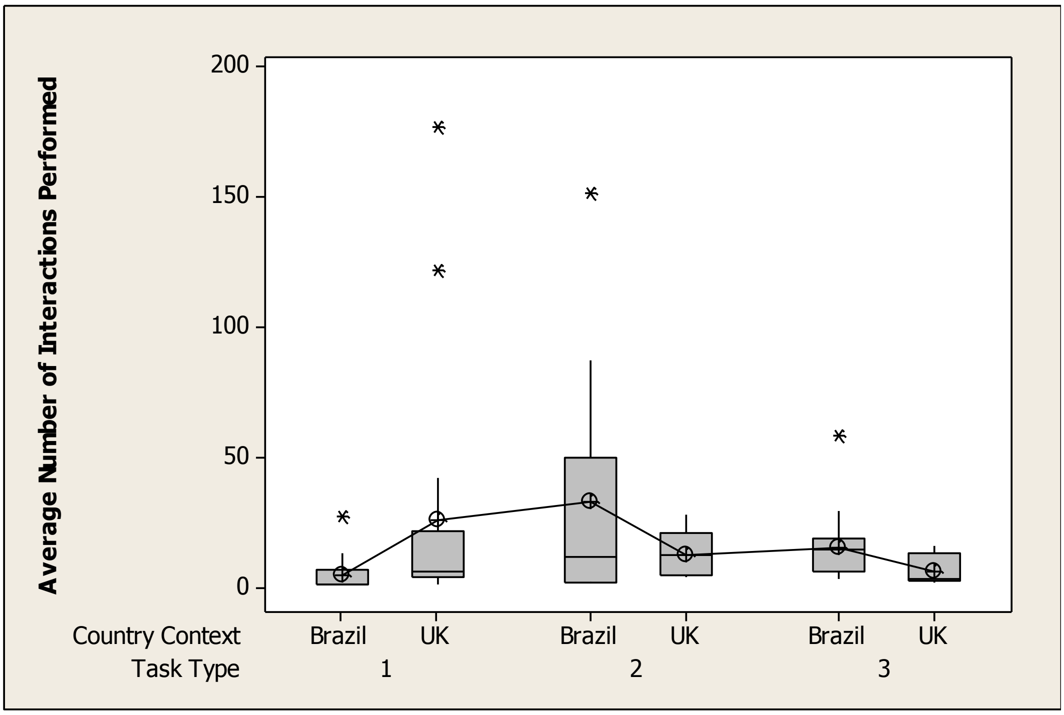


Figure 71. Average number of interactions by country context and task type

There was not enough data to provide evidence that the number of tasks accomplished by each task type, the time taken to accomplish tasks, the number of hints given to participants, and the number of interactions performed was statistically significant different for the four conditions participants were submitted.

Participants were asked to accomplish one task type using each one of the input devices, one task with the mouse and one task with the remote. There is not enough evidence that there is a statistically significant difference among the number of tasks accomplished with each device for tasks type one, two or three. Nevertheless, it can be seen from the Figure 72 below that for tasks one in which participants had to select a video already in display, more tasks on average were accomplished using the remote control than using the mouse. For tasks of the type two in which participants had to select a category in order to find a video also there were more tasks accomplished using the remote than using the mouse. For tasks three, however, in which participants had to select two categories in order to find the video, more tasks were accomplished on average using the mouse than using the remote control.

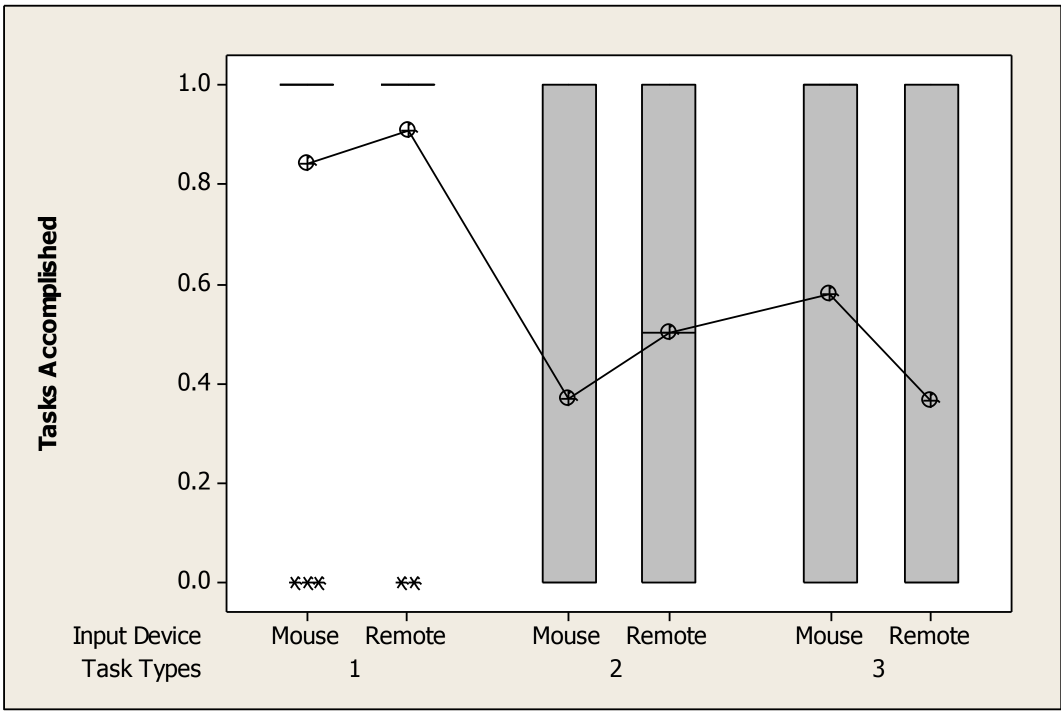


Figure 72. Accomplished tasks on average by task type and input device

Regarding the time to accomplish tasks, it can be noticed from the Figure 73 below that participants accomplished task two on average faster using the mouse than using the remote, while for tasks one and three the average time taken with each device was similar.

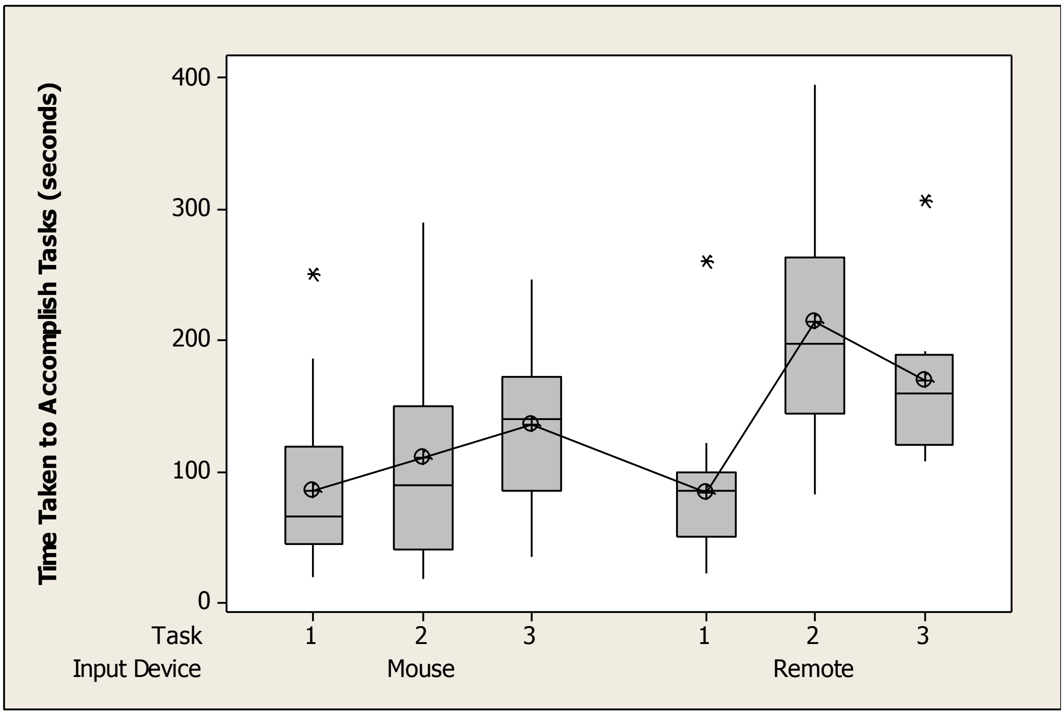


Figure 73. Average time to accomplish tasks by input device and task type

Overall more hints were given for the navigation using the remote than using the mouse, and the difference increases with the increase on task complexity. For task two in which participants had to select a category the hints for mouse navigation are on average (mean) 0.14 compared to an average of 2.8 hints to navigate using the remote for the same type of task (Figure 74).

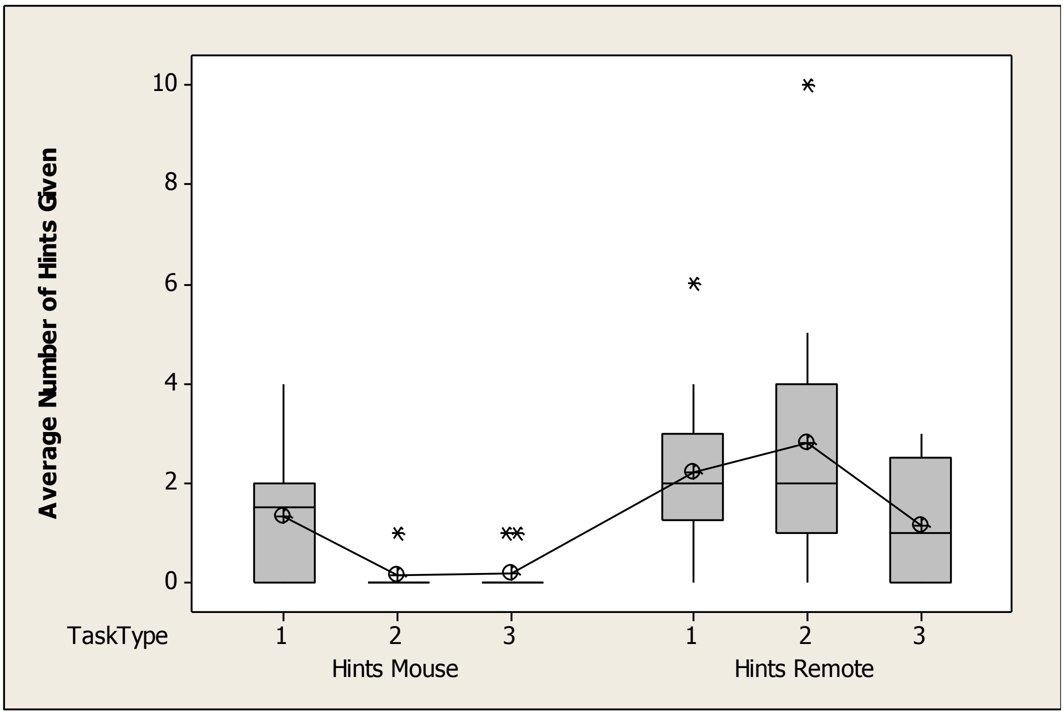


Figure 74. Average hints given by task type and input device

For the three tasks types on average (median) a larger number of interactions were performed to accomplish tasks using the remote control compared with the number of interactions performed for tasks accomplished with the mouse (Figure 75).

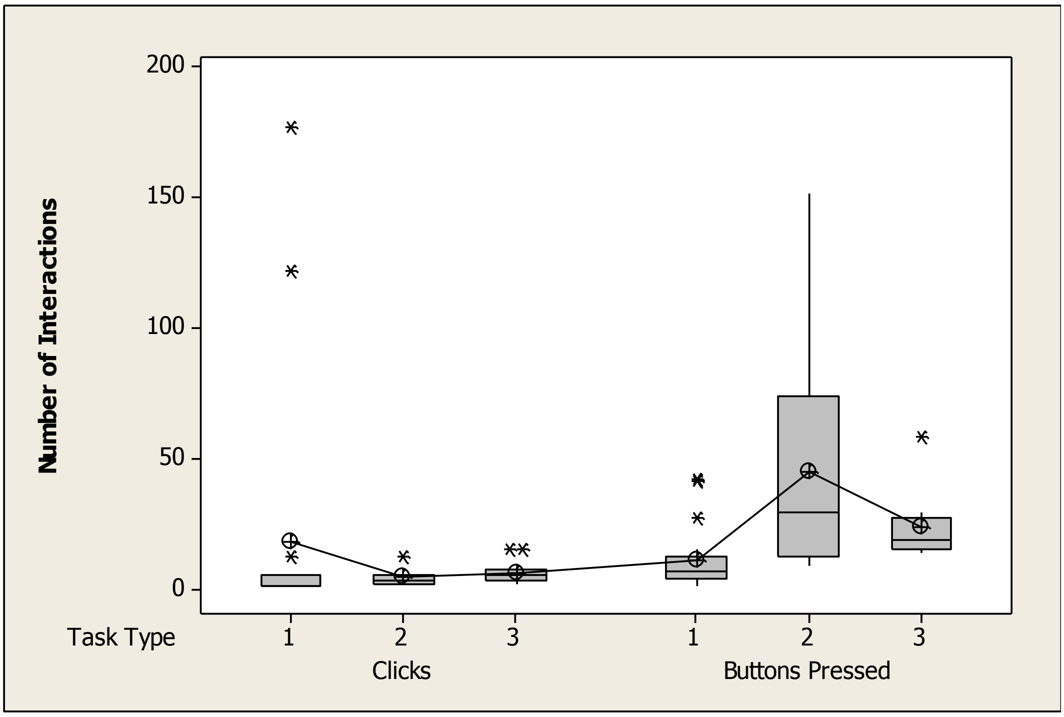


Figure 75. Average number of interactions by task type and input device

5.3.3 Participants’ Preferences

Participants’ preferences were measured in two ways: their choice of device for the free choice video search and the survey responses (Fun Toolkit). The free choice video search took place after the six tasks, when participants were told they could choose a video of their own choice to watch if they wanted, and were given the option to choose the device they would like to use. The Fun Toolkit included adapted versions of the Smileyometer, Again-Again Table and Ranking System to measure participants’ device preference. Data regarding participants’ preferences, gathered using the two methods described above, are reported in this section followed by a brief analysis of these preferences considering: participants’ age, gender, country context, condition and media and device usage.

5.3.3.1 Overview of Participants’ Preferences

5.3.3.1.1 Option of Device for the Free Choice Search

Participants’ option of device for the free choice video search varied significantly. Seven participants chose the remote control, nine participants chose the mouse and one participant chose the touch pad (Figure 76).

Among the 22 participants, five participants did not opt for any device. Two participants did not want to choose a video of their choice to watch. Three participants were not offered the option to choose, two because there was no time left, otherwise the session would exceed the thirty minute limit, and one participant was not offered to choose because he looked fidgety (Participant 18).

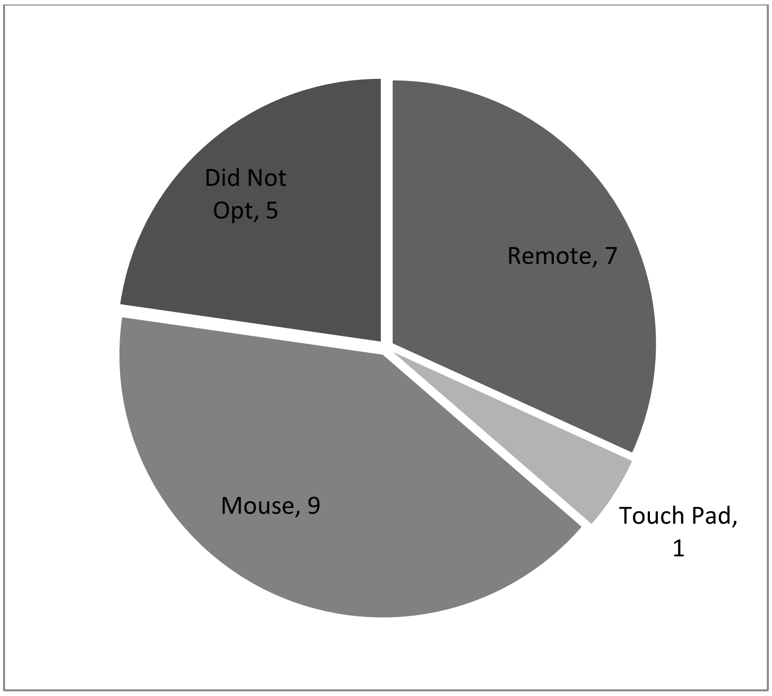


Figure 76. Number of participants who opted for each input device

5.3.5.1.2 Participants' Survey Responses (Fun Toolkit)

Participants' survey responses for the Smileyometer could vary from Brilliant to Awful; the number of participants who chose each sticker representing their opinion about each device is reported below (Table 16).

Smileyometer Responses	Remote Control (number of participants)	Mouse (number of participants)
Brilliant	9	6
Really Good	4	5
Good	2	6
Not Very Good	5	0
Awful	2	5

Table 16. Participants' Smileyometer responses

In the Again-Again Table participants responded if they would like to use each device again by colouring the words "yes", "maybe" or "no" underneath the picture of each device. Participants' responses are presented on the table below (Table 17). Some data gathered during this activity was invalid and discarded; Participant 8 did not colour any response for the remote control and Participant 12 coloured all three responses for the mouse and none for the remote control.

Again-Again Table Responses	Remote Control (number of participants)	Mouse (number of participants)
Yes	9	16
Maybe	5	5
No	6	0

Table 17. Participants' Again-Again Table responses

For the ranking, participants were asked to stamp a star underneath the picture of the device they thought was the best one to interact with the prototype. The number of participants who choose each device as the best one is presented below (Table 18).

Ranking	Responses (number of participants)
Remote Control	10
Mouse	12

Table 18. Participants' ranking responses

Participants' responses for the three activities in the Fun Toolkit are listed on the table below (Table 19).

Participants	Smileyometer	Again-Again Table	Best Device
Participant 1	Remote: really good Mouse: awful	Remote: maybe Mouse: maybe	mouse

Participants	Smileyometer	Again-Again Table	Best Device
Participant 2	Remote: brilliant Mouse: awful	Remote: yes Mouse: yes	remote
Participant 3	Remote: not very good Mouse: really good	Remote: yes Mouse: yes	mouse
Participant 4	Remote: brilliant Mouse: brilliant	Remote: yes Mouse: yes	remote
Participant 5	Remote: really good Mouse: awful	Remote: yes Mouse: yes	remote
Participant 6	Remote: brilliant Mouse: awful	Remote: no Mouse: yes	mouse
Participant 7	Remote: awful Mouse: really good	Remote: no Mouse: yes	mouse
Participant 8	Remote: brilliant Mouse: brilliant	Remote: invalid answer Mouse: yes	mouse
Participant 9	Remote: brilliant Mouse: good	Remote: no Mouse: yes	remote
Participant10	Remote: not very good Mouse: brilliant	Remote: yes Mouse: yes	mouse
Participant11	Remote: brilliant Mouse: good	Remote: yes Mouse: maybe	remote
Participant12	Remote: not very good Mouse: really good	Remote: invalid answer Mouse: invalid answer	remote
Participant13	Remote: not very good Mouse: really good	Remote: yes Mouse: yes	mouse
Participant14	Remote: good Mouse: brilliant	Remote: no Mouse: yes	remote

Participants	Smileyometer	Again-Again Table	Best Device
Participant15	Remote: brilliant Mouse: good	Remote: yes Mouse: yes	mouse
Participant16	Remote: not very good Mouse: good	Remote: no Mouse: yes	remote
Participant17	Remote: awful Mouse: brilliant	Remote: maybe Mouse: yes	mouse
Participant18	Remote: really good Mouse: awful	Remote: maybe Mouse: maybe	mouse
Participant19	Remote: good Mouse: good	Remote: maybe Mouse: yes	remote
Participant20	Remote: brilliant Mouse: really good	Remote: maybe Mouse: maybe	mouse
Participant21	Remote: brilliant Mouse: good	Remote: no Mouse: maybe	mouse
Participant22	Remote: really good Mouse: brilliant	Remote: yes Mouse: yes	remote

Table 19. Participants' responses for each Fun Toolkit activity

It may be noticed that participants' responses, if compared across the three different activities, present some inconsistencies. Participant 6, for instance, rated the remote as Brilliant and the mouse as Awful on the Smileyometer, for his responses on the device he would like to use again on the Again-Again table he coloured "no" for the remote and "yes" for the mouse and finally he voted the mouse as the best device. Such inconsistencies make it hard to measure participants' preferences accurately, therefore, to facilitate the process and assist calculating if preferences differ considering participants' age, gender, country context, condition and their media and device usage, participants' responses were converted into numbers. Responses were coded numerically from one to five, in which one and two were equivalent to negative responses, three was a neutral response and four and five were positives responses about each device. For the Smileyometer, the values attributed to each response were: one (awful), two (not very good), three (good), four (really good) and five (brilliant). For the Again-Again table the values were: one (no), three (maybe) and five (yes). Finally, for the Ranking: one (device other than the best) and five (best device). These numbers were then added into mouse and remote responses for each participant resulting on a value between three and 15 that reflects participants' degree of preference for each device respectively (Table 20).

Participants	Remote Control (degree of preference)	Mouse (degree of preference)
Participant 1	8	9
Participant 2	15	7
Participant 3	8	14
Participant 4	15	11
Participant 5	14	7
Participant 6	7	11
Participant 7	3	14
Participant 8	6	15
Participant 9	11	9
Participant 10	8	15
Participant 11	15	7
Participant 12	7	5

Participants	Remote Control (degree of preference)	Mouse (degree of preference)
Participant 13	8	14
Participant 14	9	11
Participant 15	11	13
Participant 16	8	9
Participant 17	5	15
Participant 18	8	9
Participant 19	11	9
Participant 20	9	12
Participant 21	7	11
Participant 22	14	11

Table 20. Participants' degree of preference for each input device

The degree of preference for the input devices is similar, but on average, mean and median, the degree of preference for the mouse is slightly higher than for the remote (Figure 77).

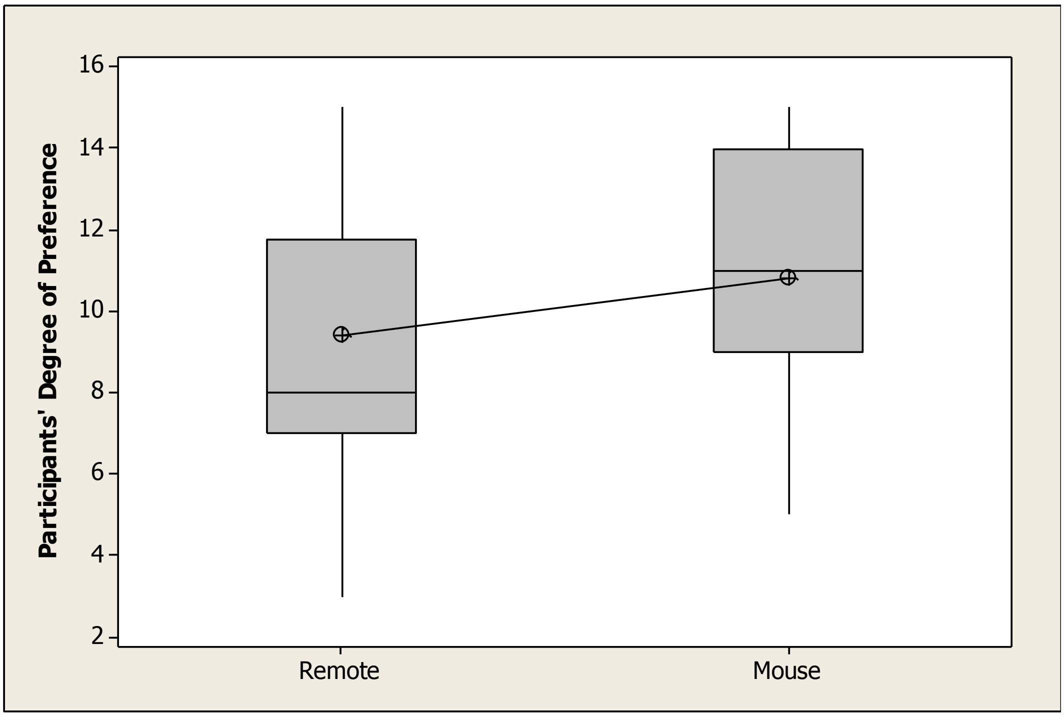


Figure 77. Participants' degree of preference for each input device

The three participants who during the evaluation sessions used an alternative input device instead of the mouse were also asked to rate the mouse in the Fun Toolkit. Since participants were first asked to interact using the mouse, it was decided to collect their opinion about the device despite the fact that they struggled to use the mouse (Participants 12 and 16) and were then given the option to use an alternative device or chose not to use the mouse (Participant 19).

Participants' option of device for the free choice video search was not included on the calculations of participants' degree of preference because a significant amount of data was missing. Data from six out of the twenty two participants were lacking or invalid, five participants did not choose any device and one participant chose an alternative input device.

5.3.3.2 Preferences and Age

There is no evidence of a statistically significant difference on the degree of preference for each device within each age group or among the two age groups. It may be noticed, however, that on average, mean and median, participants' preference for both age groups is higher for the mouse than for the remote. In addition, it can be seen that four years old participants' degree of preferences is higher than three years old participants' for both devices (Figure 78).

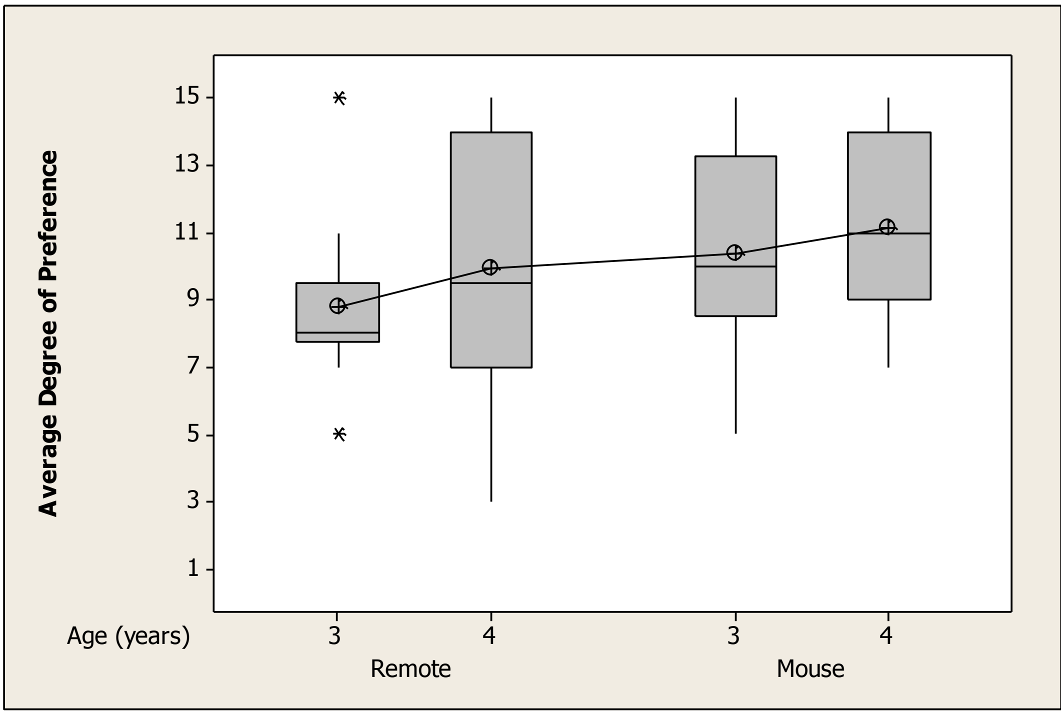


Figure 78. Average degree of preference for each input device by age

5.3.3.3 Preferences and Gender

There is no evidence of a statistically significant difference on the degree of preference for each device for boys or girls. The preferences between the two genders are also similar for the remote and for the mouse.

5.3.3.4 Preferences and Country Context

There is no evidence of a statistically significant difference on the median of the degree of participants' preferences for each device for participants in Brazil and in the UK. It may be noticed that the average (mean and median) degree of preference for participants in both countries are slightly higher for the mouse than for the remote. The median for both devices for both country contexts is exactly the same; there is however a difference on the mean that suggests that degree of preference for the remote for participants in Brazil is higher than for participants in the UK (Figure 79).

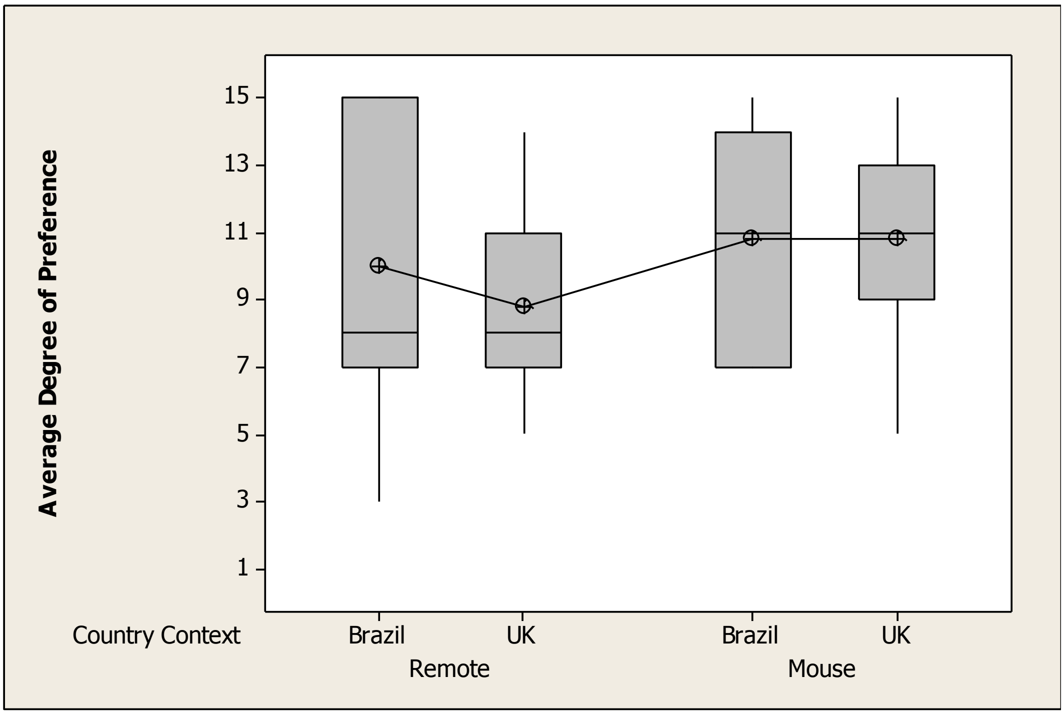


Figure 79. Average degree of preference for each input device by country context

5.3.3.5 Preferences and Condition

There is no evidence of a statistically significant difference on the median of the degree of participants' preferences for each one of the four conditions for the remote or for the mouse.

5.3.3.6 Preferences and Participants' Media and Device Usage

There is no evidence of a statistically significant difference on the median of the degree of participants' remote control preferences for those participants who watch television several times per day, about once a day or a couple of times a week.

There is no evidence of a statistically significant difference either on the degree of participants' mouse preferences for participants who use the computer more frequently, on a daily or weekly basis, compared with those whose parents' declared use the computer occasionally, on a monthly basis or less often.

Considering the remote usage, there is no statistically significant difference on the degree of preference for the remote control for participants who use the remote, those who do not use it and participants who only use the device with supervision.

There is, however, a difference on the degree of preference for the mouse for participants who use the mouse, those who use it only with supervision and participants who do not use the mouse. The average (mean and median) degree of preference for the mouse of participants who use the mouse is significantly higher than that for participants who do not use the mouse or use it only with supervision (Figure 80).

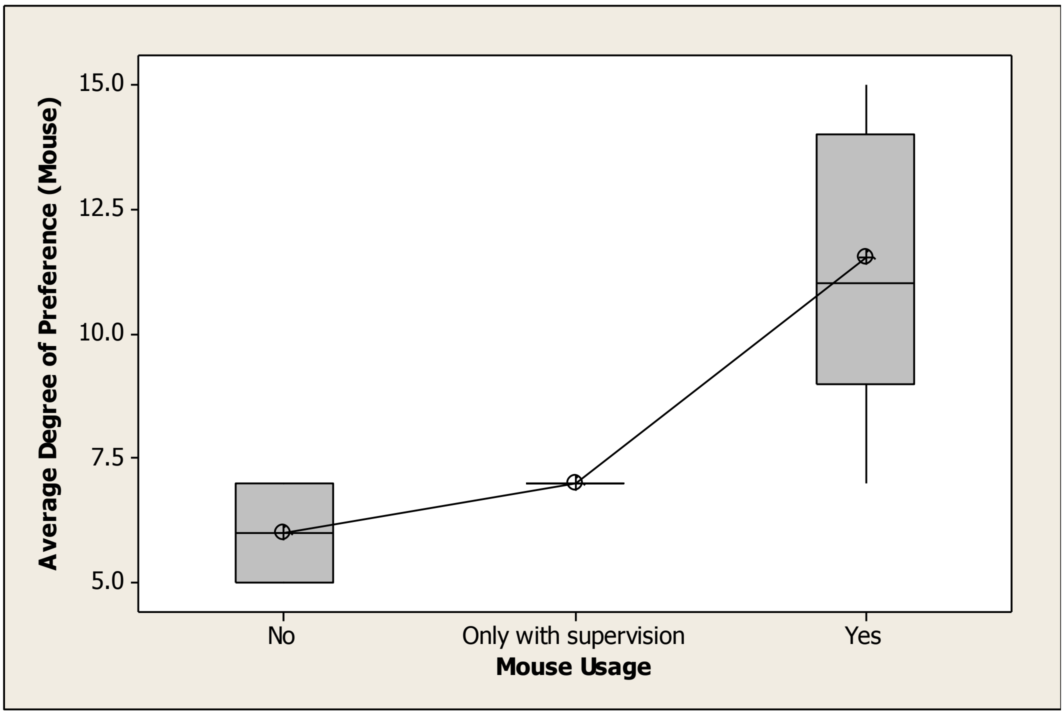


Figure 80. Average degree of preference for the mouse by mouse usage

5.3.3.7 Preferences and Tasks Accomplished

There is no evidence of a statistically significant difference on the degree of preference for each device and the number of tasks participants accomplished using the device. It may, however, be noticed that the degree of preference for the remote is on average (mean and median) lower for participants who did not accomplish any task with the device (Figure 81). The degree of preference for the mouse is on average (mean and median) very similar for participants who accomplished one, two or three tasks using the device, but it is slightly lower for those participants who did not accomplish any tasks using the mouse (Figure 82).

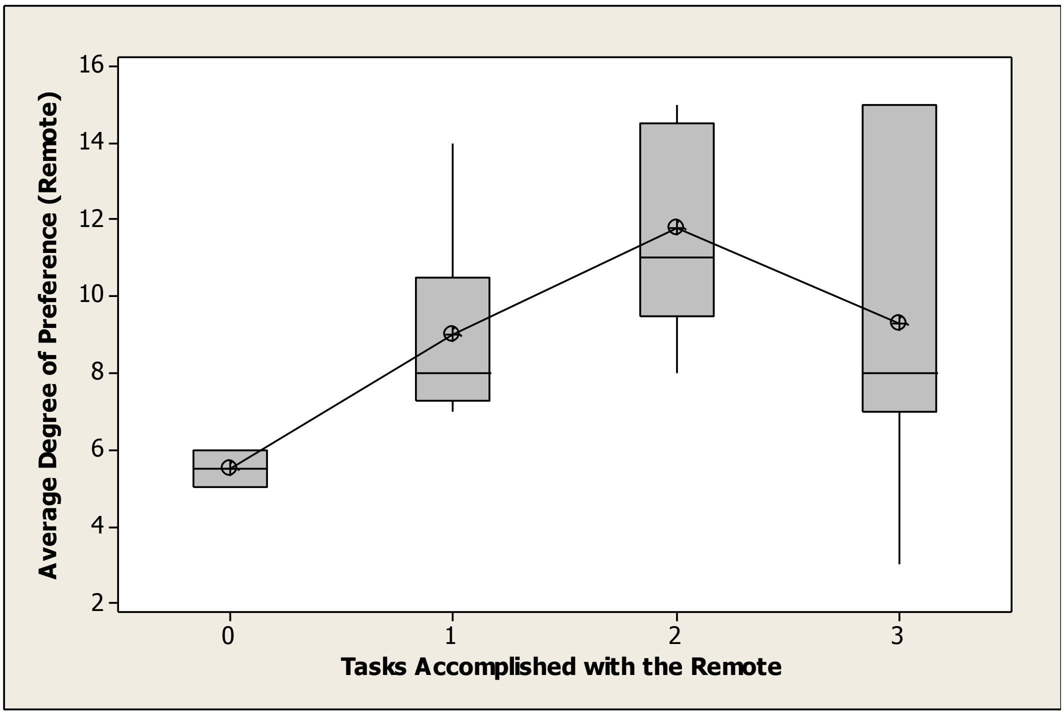


Figure 81. Degree of preference for the remote by tasks accomplished with it

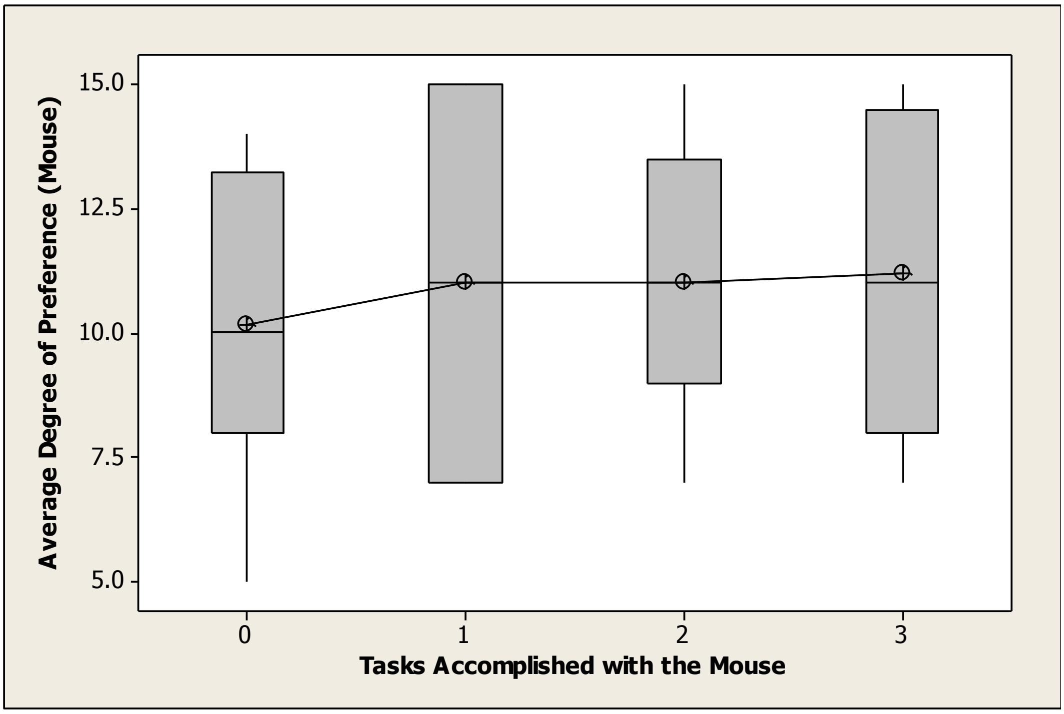


Figure 82. Degree of preference for the mouse by tasks accomplished with it

It seems that the degree of preference for an input device is lower if participants did not accomplish any tasks with the device, but there is not enough evidence that the degree of preference increases accordingly as participants accomplish one or more tasks of different types and complexity with a device.

5.3.3.8 Preferences and Time Taken to Accomplish Tasks

There is no evidence of a statistically significant difference on the degree of preference for the remote or the mouse for participants who took a long or short time to accomplish tasks with each device respectively.

5.3.3.9 Preferences and Hints Given

There is no evidence of a statistically significant difference on the degree of preference for the remote or the mouse for participants who were given a larger or smaller number of hints to accomplish tasks using each one of the devices.

5.3.3.10 Preferences and Interactions

There was not enough evidence that the degree of preference for each device was different, according to the average number of interactions performed to accomplish tasks for the remote or the mouse.

The median of the degree of preference for the remote was similar for participants who pressed a large or small number of buttons to accomplished tasks, as the median of the degree of preference for the mouse was similar for participants who performed a large or small number of clicks to accomplished tasks.

5.3.4 Participants' Interaction with the Prototype

In the previous sections, the data collected during the evaluation sessions was analysed statistically, illustrated with graphs, tables and, when appropriate, compared with some findings from the video analysis. Thus far, most data presented was numerical. There was, however, a lot of qualitative data collected that can provide further details and elucidate the process of children's interaction with iTV applications. In this section, the qualitative data will be explored in more detail. It will be reported how participants' physical and cognitive abilities interfered with their interaction with the prototype, the way participants talked about their understanding of the interface and how the interaction with the prototype worked as well as some comments that reflect participants' needs and preferences.

5.3.4.1 Participants' Physical and Cognitive Characteristics

Children's handedness was the first physical characteristic considered an issue for the interaction process. Participants were prompted to use the remote control and mouse with the hand they felt most comfortable using. For the interaction using the remote control, most participants used both hands. For the interaction with the mouse, however, when told to use the device with the hand they felt comfortable using, several participants looked puzzled and the question had to be complemented by asking which hand they used to draw. Eventually, most participants chose to use the mouse on the right side but there were two participants who chose to use it on their left.

Some participants had found holding the remote control difficult. Two participants (Participant 1 and 2) told the researcher the remote was slippery. After each session, however, it was confirmed that there was no grease or anything on the device itself that could be causing the remote to slip. Most participants rested the remote on the black piece of plastic over the keyboard and touch pad. For 12 out of the 22 participants, resting the remote on the keyboard interfered with the content being displayed on the screen, by accidently moving the cursor via the touch pad. As reported previously (Section 5.2.1), some additional participants accidently pressed keys that were equivalent to short cuts which closed the recording programmes running behind the prototype, and as a result some data was lost and they had to be excluded from the study. Overall it was observed during the sessions that participants in general held the remote with both hands and rested it over the keyboard and touch pad to press the buttons. The webcam, however, recorded only participants' faces; details of their hand movements were not captured so could not be analysed further.

Hand-eye coordination issues were also observed during the evaluation sessions in both younger and older participants. Several participants struggled when trying to stop the cursor on target objects such as screenshots and buttons using the mouse. For the interaction with the remote control, the cursor movements were broken into discrete steps in which each press of an arrow button moved the cursor one step. Although the cursor stopped on each possible target, some participants also found it hard to stop on the target object they wanted using the remote control because they pressed the buttons repeatedly. To select target objects, access videos through screenshots or turn buttons on and off, some participants focused more on the screen to aim at their target and then looked at the mouse to click. Other participants looked only at the screen. Using the remote control, some participants looked at the remote to press the buttons then back to the screen then to the remote to press another button and so on. Other participants only looked at the remote for the first time they were pressing a button, then navigated through the options and selected looking only at the screen.

As previously mentioned (Section 5.3.2.6), two three year old participants (Participants 12 and 16) were offered the use of an alternative input device because they struggled to interact using the mouse. Both were told to use the touch pad, but when they tried to use it, it did not make the interaction easier, so it was suggested they use the keyboard. Participant 12 tried three times to interact by directly touching the screenshots she wanted to select on the computer screen and had to be told that it was not a touch screen interface.

5.3.4.2 Participants' Understanding of the Interface

This section includes some participants' actions and comments that may elucidate how different elements of the prototype application were understood, misunderstood or not understood.

5.3.4.2.1 Screenshots

During the evaluation sessions it was found that, as discussed in Chapter 4 (RCS2), most children were able to relate screenshots to the video content they represented. The videos participants were asked to find were well known, so most participants were familiar with most videos. It was noted, however, that some participants told the researcher they were familiar with the video when first asked to find it but as the screenshot was displayed on the screen they did not recognise it. The researcher let the screenshot pass twice on the 3D wheel loop and in case the participant did not recognise it, s/he was then shown the image and description of the video as shown to participants who had said they were not familiar with the video to be found. Other participants that affirmed being familiar with the video they were asked to find selected a screenshot equivalent to another video instead.

In total 20 videos accessed by 14 participants were not the videos they were requested to find, and in these cases participants failed to accomplish a task by selecting a screenshot other than the one from the video to be found. It can be noticed, however, that most of the screenshots selected were strong members of the same category as the screenshot to be selected. The screenshot from the video

Cinderella was the one that most participants failed to recognise, with ten participants selecting a screenshot from another fairy tale instead.

There were, on the other hand, screenshots recognised by a significant number of participants. Ten participants in Brazil recognised and selected the screenshot from the television programme *Cocoricó*. One participant who was not familiar with the programme had to be shown the image, after which they accessed the correct video. Nine participants in the UK recognised the *Pingu* screenshot, and two participants had to be shown the image of the cartoon, but all 11 participants selected the screenshot correctly and accessed the *Pingu* video completing the task.

Screenshots that were not part of tasks to be accomplished were also recognised and provide evidence that most participants clearly understood that screenshots were related to the video content they represented. Participants recognised screenshots being displayed on the 3D wheel and mentioned videos they had, *this is Alvin and the Chipmunks, I have their DVD* (Participant 1); videos they like, *The Incredibles I saw, yeah, that's my one, it is my favourite cartoon* (Participant 13); and videos they watch on TV, *Mister Maker, hey, I watch it on the Discovery* (Participant 3) referring to the programme broadcast on the Discovery Kids cable channel in Brazil.

The most impressive screenshot recognition was made by Participant 9. For the free choice task she mentioned she owned and would like to watch *the time machine one;* she then accessed a Brazilian cartoon through the screenshot displayed below (Figure 83). The screenshot is an excerpt of the film *A Time Adventure* by the cartoonist Maurício de Sousa in which one of the characters, who does not feature in the screenshot, builds a time machine. It was impressive that the participant was able to recognise it because the screenshot does not contain anything related to the time machine itself, and there are a series of short and feature length films with those same characters that were shown in the cinemas, are available on DVD, for download on the internet and broadcast on a weekly basis on a cable television channel in Brazil.

Figure 83. Screenshot from the animated Brazilian film *A Time Adventure*

5.3.4.2.2 Buttons, Categories and Icons

Participants were asked to accomplish six tasks in total, two of each type. The tasks of the first type did not require a category to be selected. The two type two tasks involved a simple search; in one they had to find a video which was a member of the 'fairy tales' category and in the other they were required to find a video which was a member of the 'superheroes' category. The type three tasks were two conjunctive Boolean searches; in one participants had to select the categories 'TV shows' and 'music and songs' and in the other they had to select the categories 'movies and films' and 'animals and nature'.

Not all the categories and icons were recognised by participants. As mentioned in section 5.3.1.3 a total of fifty eight hints were given to participants to assist them in finding a video by selecting the appropriate category. Among these hints, forty eight were hints about the category that assisted in performing a simple (H10) or Boolean search (H11), and ten were hints given to participants who selected the wrong category (H12).

When necessary, participants were told the categories to which the video to be found belonged, and in case no action was taken and/or they needed further assistance participants were told to look for the icon representing the category on the buttons on the bottom of the screen (H9). There were, however, no hints about the icons.

When told to which category the video belonged and asked where they thought it could be found, some participants demonstrated they did not recognise the category icon. When asked where she thought fairy tales were, Participant 10 responded *I don't know*. Participant 21's answer for the same question was *I don't know which one*, meaning he did not know which button referred to fairy tales. Participant 22 also responded *I don't know* for the questions about where the 'fairy tales', 'superheroes' and 'music and songs' could be found. Participant 6 when asked where 'animals and nature' could be found said *there are no animals here*. Additionally, when asked where he could find music, Participant 2 said *in this fish one* referring to the 'animals and nature' button.

Some participants did not answer the hint question or talked about their search, so they did not demonstrate the icon for a category was not recognised, but navigated using the mouse or the remote over several buttons until they found the category they were looking for by hearing the category label through the audio feedback.

Some icons were misunderstood by participants. Participant 6 selected the button for the Nick Jr. channel, whose icon is the channel's logo represented by an elephant, when asked where he thought the category 'animals and nature' was. Participant 10, when told to find the film *Ratatouille*, part of the same task, also selected the Nick Jr. button, in her case before any hints were given.

There were icons that were recognised but misinterpreted. Participant 1, for instance, was given hint 11 to select two categories and perform a Boolean search. He was told to find a music video from a TV show and asked where he thought he could find ‘music and songs’. He navigated through the buttons and turned the ‘music and songs’ button on. Then he was asked where he thought he could find ‘TV shows’. He looked at the researcher and asked *where? In this one with the computer?* referring to the TV shows icon represented by a television screen.

There were some videos that participants identified as members of a category to which they belonged, but were not strong members of. Two participants (Participants 14 and 15) identified the fairy tale Cinderella as belonging to the ‘cartoons and animations’ category that it actually did; and two participants (Participants 1 and 19) thought Cinderella would be a member of the Disney category and in this case it was not because the Disney button on the prototype included only television programmes and series broadcast on the channel Playhouse Disney, not all films produced by Disney. The film Ratatouille was also recognised by two participants (Participants 1 and 20) as being a member of the Disney category, and like Cinderella, Ratatouille was produced by Disney but not included on the Disney button on the prototype application. Two other broadcasters were recognised as categories. Three participants (Participants 1, 3 and 18), two in Brazil and one in the UK, recognised the superhero cartoon Ben 10 as being a member of the Cartoon Network category, the channel on which it is broadcast in Brazil and United Kingdom. Two participants in the UK (Participants 13 and 17) recognised Pingu and Tweenies respectively as members of the CBeebies category, the channel on which both programmes are broadcast in the United Kingdom. In all those cases participants recognised both the category and the channels’ logo represented as icons on the prototype’s buttons.

Participants demonstrated that they recognised several other categories and icons on the prototype. When participants pointed at icons on the screen no feedback was given if the icon did not represent the category of the video they were looking for, but in cases where they pointed at the correct icons, the researcher complemented them, because through the video captured via the webcam it would not be possible to identify to which icon participants were pointing at during the data analysis. This was unexpected, and was not included during the planning of the evaluation sessions, but it was an improvised way the researcher found to register the data.

When asked where they thought fairy tales could be found (H10), four participants (Participants 4, 11, 13 and 15) pointed at the ‘fairy tales’ button on the screen. Participant 9, when asked where she could find *Cinderella*, said *in the wand one*, with no category hints; Participant 20 also when asked where she could find *Cinderella* said without category hints *in the one with the wand on it*.

Regarding the ‘superheroes’ category, Participant 5, when asked where she thought superheroes were (H10), pointed at the ‘superheroes’ button. When asked where she could find *Ben 10* with no category hints, Participant 19 said *the boyish one*, and when asked which one she thought was the boyish one she pointed at the ‘superheroes’ button on the screen. Participant 6 and Participant 9, when asked to find the *Ben 10* video, navigated with the mouse directly to the superheroes button, turned it on then navigated to the upper screen and clicked on the *Ben 10* screenshot displayed; in seventeen and twenty eight seconds respectively, the two participants recognised the category to which the video belonged and the icon that represented it and accomplished the task with no hints.

The ‘music and songs’ icon was recognised by three participants. When asked where she thought music could be found, Participant 9 pointed at the ‘music and songs’ button. Participants 17 and 19, when asked where they thought they could find music and songs, navigated directly to the ‘music and songs’ icon and turned the button on.

The ‘TV shows’ icon was also recognised by Participant 10. When asked where she thought TV shows could be found (H11), she pointed at the ‘TV shows’ button on the screen.

With regard to the ‘movies and films’ category, Participant 1 pointed at the ‘movies and films’ button and said *here the film* when asked where he thought films were (H11). When, during the same task, Participant 9 was asked where she could find films, she navigated using the mouse directly to the ‘movies and films’ button and turned it on.

Some participants also recognised the ‘animals and nature’ icon. When asked where he thought he could find animals and nature (H11), Participant 11 pointed at the ‘animals and nature’ button on the screen. When asked the same question about the ‘animals and nature’ category (H11), Participant 9 navigated using the mouse directly to the ‘animals and nature’ button on the screen and turned it on.

There were some participants that needed hints to recognise the buttons in which icons did not represent categories, but instead the functions back and clear all. Five hints in total were given to assist on the navigation through these buttons and to help participants accomplish tasks. One hint to use the clear all button (H7) was given to Participant 2 who selected too many categories and would have had to turn off each button individually using the remote. Three hints were given to use the back button and return to the menu, one using the mouse (H4) and two using the remote (H8). Participants 2 and 14 were given the hint (H8) because it was thought they had accessed a wrong video accidentally. Participant 2 pressed OK while trying to navigate to the buttons on the bottom of the screen and Participant 14 pressed OK almost immediately when asked to accomplish the first task, accessing the first screenshot selected without navigating. Moreover, Participant 18 was given the hint (H4) to return to the menu because he accessed a video before considering the task he was just asked to perform.

These were, however, only hints concerning the coloured functional buttons given to assist the participant in accomplishing tasks. As participants accessed the video segments, they were told they could go back to the menu if they wished and then some recognised the back icon and navigated on their own while others needed a hint (H4 or H8). This was also something that was unpredicted. During the first session it was found that the videos were short but the participants could still become bored watching over two minutes of a programme they did not enjoy and this could later be reflected in their experience with the prototype and influence their opinion about the device being used to interact at that point. For this reason, it was decided to tell participants they could return to the menu if they wished after accessing the video they were asked to during the task. Some participants decided to watch the video segments until the end, so did not use the back button. Two participants, one in Brazil (Participant 3) and one in the UK (Participant 22), recognised and used the back button with no hints when told they could go back to the menu. The other 20 participants were given a hint (H4 or H8) to go back on the first occasion. Five participants among those used the back button with no hints during the following tasks. Five needed an additional hint (H4 or H8) during the next task. Two needed additional hints (H4 or H8) for the two following tasks and eight participants chose not to use the back button in the following tasks.

No hints were given to help participants in recognising the icons for the close, favourites and help functions because they were not part of the process to accomplish tasks. There were, however, participants who used these prototype features during the interaction and some issues with icon recognition are worth reporting.

Participant 1 misinterpreted the red close button. When asked to find a music video from a Brazilian television programme whose logo is a red “X”, he pointed at the close button on the prototype and said *it’ll be here in the little red*. He was then told to roll the mouse over the button to confirm it was not the category he was looking for. During the same Boolean search, Participant 2 selected a button other than those from the two categories he was supposed to select. He was asked if that was correct (H12) then asked if he knew how he could turn the button off and deselect the category. He said *I know the red* and pressed the red button on the remote and the prototype closed. He was then told to press any button to reopen it and chose to press the green button to open the application. It is not clear whether he related the icon on the red button to the TV programme's logo, he associated the red button with turning off the selected category or he really wanted to close the prototype. When looking for the same music video, Participant 8 selected the 'music and songs' button then the 'TV shows' button, and then recognised and pointed at the screenshot of the music video to be found on the screen. However, as she moved the cursor from the buttons to the upper screen, the screenshot had disappeared on the 3D wheel and instead of waiting for it to reappear and click on the screenshot she rolled over the close button twice and closed the prototype. In this case the participant did not speak a word during the interaction, so it is not evident whether she related the close button to the logo of the video to be found and so by clicking on it expected that the screenshot would return to the visible part of the wheel, or if she actually meant to close the prototype.

Participant 13, on the other hand, clearly closed the prototype on purpose. She interacted using the remote first and during the first task with the mouse when asked to find *Dora the Explorer*, she rolled over the close button twice, heard the audio feedback and clicked around the button several times until she managed to click on the close button and closed the prototype. When told to click again to reopen the application she said *I don't want to open*. She said she did not want to open it because she wanted to watch something else instead of the *Dora* video she was asked to find, as previously mentioned (Section 5.3.2.3). During the following task, the same participant (Participant 13), when asked to find the superhero cartoon *Ben 10*, said *Ben 10? Let's go back to…* She did not complete the sentence, but navigated using the mouse directly to the close button and closed the prototype once more.

Participant 15 closed the prototype accidentally. During the first task using the remote he pressed the OK button three times so that within a few seconds of the video being accessed, he returned to the menu and then accessed the video again. As he accessed the video for the second time he was told he could go back to the menu if he wanted, after which he pressed the OK button one more time to go back. However, as the video had just finished playing, instead of going back to the menu, the application was closed. He pressed the OK button repeatedly three times; opening, closing and reopening the prototype in the process.

The favourites feature was not included in the evaluation sessions, but a few participants used it. Participant 8 accessed the cartoon *Charlie and Lola* instead of the video she was asked to find for the first task using the remote. The screenshot for the *Charlie and Lola* cartoon was next to the screenshot of the *Dora the Explorer* cartoon she was required to find and it was not evident whether she accessed the other video accidentally or on purpose. As she watched the video she was told she could go back to the menu if she wanted and was given the hint to navigate using the remote (H8). She did not say anything, but pressed the yellow button instead and made the video her favourite. She then pressed the red button and went back to the menu. For the final choice task, Participant 8 chose from the 3D wheel the same *Charlie and Lola* screenshot to watch the video again. When the video was over, while the researcher got the Fun Toolkit form and materials for her to fill out the survey, she navigated using the mouse to the ‘favourites’ button, turned it on and off a couple of times and observed the *Charlie and Lola* screenshot presented on the screen as the button was selected.

Participant 14 was thought to have made a video favourite by accident. He was watching a video segment other than the one he was asked to find for the simple search task using the mouse; he was told he could go back to the menu if he wanted and clicked on the yellow ‘favourites’ button instead of the back button. He clicked on the button twice, turned it on and off, then clicked on the help button and accessed the help section. After accomplishing the first task using the remote, Participant 14 was watching the video he was asked to find and a similar issue occurred. He was told he could go back to the menu, and instead of pressing the red button he pressed a series of invalid buttons such as numbers, then the yellow button, then the blue button and accessed the help section one more time.

While watching a video, Participant 5 was told she could go back to the menu, and went through a similar process. She clicked on the yellow ‘favourites’ button, turned it on and off, then clicked on the help button and accessed the help section. During a previous task she had already accessed the help section while trying to go back to the menu. She was watching a video segment and was told she could go back if she wanted and said *I’ll go back*, then rolled over the buttons and clicked on the help button instead of on the back button.

Five additional participants also accessed the help section during their interaction with the prototype. Participant 1 accessed the help section when asked to find the fairy tale *Cinderella*. He navigated to the help button, rolled the mouse over it twice, heard the audio feedback, aimed and clicked, thus accessing the help section. When asked to find ‘music and songs’ (H11), part of a Boolean search, he also navigated to the help button and clicked on it accessing the help section one more time.

Participant 2 also accessed the help section when trying to accomplish a Boolean search. He rolled the mouse over the help button and as he heard the audio feedback he repeated *help?* then rolled the mouse over the help button two more times, clicked and accessed the help section.

Participant 15 accessed the help section when asked to find the *Dora the Explorer* cartoon, already on display on the 3D wheel. He rolled over the help button four times, heard the audio feedback and clicked, accessing the help section.

Some participants tried to interact with the help section. Participant 5 clicked twice on the red back button on the tutorial instead of clicking on the button on the actual prototype. Participant 6 accessed the help section while trying to perform a Boolean search; he clicked on the help button twice, accessing segments one and two, but did not watch either until the end. In addition, during the first help segment video, he mimicked the cursor movements on the screen with the mouse. It also seemed like he was trying to interact with the non-interactive version of the prototype being displayed on the screen to illustrate how to interact. Participant 17 used the buttons on the actual prototype, interacted and interfered with the tutorial presented; she clicked on the help button several times during the tutorial, accessing the two different parts of the help section, and did not watch them until the end either. The researcher asked her to watch the tutorial, and presented it a second time, but the same occurred. She clicked on the buttons and interacted, going back to the menu one more time.

Participants did not talk about the help icon, so there is not enough evidence that they recognised the icon before hearing the audio feedback revealing the function of the button, and it is not clear what type of assistance they expected to find through its use.

The pause icon displayed on the bottom left of the screen was not recognised by any participant. Some participants (Participants 2, 5, 7) did not respond when asked if they thought they could pause the video if they wanted. Two participants related pausing to going back to the menu when asked if they knew how to pause the video; Participant 6 said *it's easy to go back* and pressed the red button and Participant 9 said *yes by pressing the red button*. The other participants were not asked the question because they did not accomplish the last task in which the question was asked.

Participant 12 was the only participant who paused a video; she was navigating through the buttons using the remote control while watching a video segment of her choice during the free choice task. As she got to the pause button, she pressed OK several times, making the video pause and play eight times consecutively. The researcher showed her where the cursor was on the screen but she continued to pause and play the video. From her facial expressions it was not evident whether she was doing that on purpose and enjoying pausing the video or was pausing the video accidentally by repeatedly pressing the same OK button.

There were some issues with the visual feedback provided indicating that the buttons were selected and/or on. Four participants (Participants 5, 6, 20 and 22) turned buttons on and off several times and seemed not to understand when the buttons were on or off.

During the evaluation sessions, no one demonstrated an understanding that the greyed buttons were inactive, but only two participants (Participants 10 and 17) tried to interact with greyed buttons by clicking on them.

5.3.4.2.3 The Interaction Process and Beyond

In this sub-section are presented additional details of participants' actions and comments, before, during and after the interaction with the prototype.

At the beginning of the evaluation session, before the tutorial was shown and any interaction started, two participants commented on the prototype. As the application was loaded Participant 5 mentioned *I have a computer like this one* and Participant 13 said *Wow!*

After watching the tutorial explaining how the interaction with the prototype worked, some participants understood the interaction process and accomplished the first tasks using each device with no hints or assistance.

Participant 2 interacted with the remote first (condition two) and needed hints to accomplish tasks, but after being shown the mouse tutorial he was asked to find the cartoon *Dora the Explorer* and said *I know how this works (...) when I see it I'll click*, he clicked on the screenshot and accomplished the first task with no hints.

Participant 3 interacted first with the mouse (condition three) and was given some hints to accomplish the tasks, but following the remote tutorial he was asked if he could find the Brazilian programme *Cocoricó* and said *there* pointing at the screenshot. He was then asked how he could watch the programme using the remote and stated *we have to go back, how can I go back?* He then looked at the remote and pointed at the arrow key *there?* He used the arrow button to navigate through the screenshots and pressed the OK button accessing the video and accomplishing the first task with no hints.

Participants 6 and 9 interacted with the remote first (condition two) and after being shown the mouse tutorial accomplished the first two tasks using the mouse with no hints. Participant 9 also accomplished the first task using the remote with no hints.

Participants 7 and 18 interacted with the remote first (conditions four and two) and following the mouse tutorial accomplished the first task using the mouse with no hints. Participant 17 also accomplished the first task with the mouse with no hints but had not interacted with the remote previously (condition one).

Some participants talked about the interaction process and their words may elucidate how they understood the interaction to work. Participant 1, during the interaction with the mouse when asked if he thought he could find *Cinderella* to watch, said *I have to press the little button when she* [*Cinderella*] *sees me*.

Participant 2 was navigating with the mouse trying to find the 'animals and nature' button. After being given a category hint (H11) he mentioned *I'm touching it to see,* referring to the process of rolling the mouse over the buttons and hearing the category name through the audio feedback.

Participant 3, while performing a Boolean search, selected one category then was given a hint (H11) to select the other category but he did not do so. He looked at the wheel and went over the screenshots presented, and did not find the one he was looking for because only one category was selected, so the screenshot for the video to be found was not presented. Then, using the mouse, he moved the cursor on the screen from the right side to the left side of the wheel and said *let's see the other side, turn, turn, hey turn to the other side* referring to the movement of the 3D wheel with the screenshots.

Other participants also commented on the wheel movement. Participant 11, while trying to accomplish the first tasks using the mouse, moved the cursor from one side of the wheel to the other and said *it has to stop, it's going too fast*. Participant 12, when asked if she could find *Dora the Explorer* to watch said *yes, it's around there* pointing at the wheel. She navigated using the remote and while trying to find *Dora*'s screenshot to accomplish this first task mentioned *oh it goes fast, it goes away from me*. Participant 13, when asked if she could find *Cinderella* to watch using the mouse, said *yes* and then asked *where stops?* referring to the wheel. Participant 19, while navigating using the touch pad looking for a screenshot she had already identified around the wheel, said *stop! a*nd Participant 1, while trying to accomplish two of the three tasks using the mouse was asked if he could find the video and said *there* and added *but it has to stay still*.

Several participants asked for additional information on the process of interaction with the prototype. Some required a clarification or more detailed information about the instructions presented in the tutorial. Participant 3, for instance, during the first task to be accomplished, moved the cursor using the mouse over the screenshot to be selected and then asked the researcher *and now? Where should I press?* While watching the tutorial for the remote, and as soon as the instructions said that the OK button should be pressed to access the videos, Participant 13 asked the researcher *where is OK?* Moments later during the same tutorial the participant recognised a screenshot and as she talked about it she missed something that was said on the tutorial, then asked the computer this time not the researcher, *can you say that one again?* She brought her ears close to the computer and asked the computer one more time *can you say again?*

Other participants asked for a confirmation they were about to press the right button or select the right category or screenshot. For example, during the first task to be accomplished with the remote, Participant 6 looked at the remote control and asked the researcher *is it OK really?* confirming the button to be pressed to access the video. Participant 10 was given a category hint (H11), navigated using the remote to the 'music and songs' button she was supposed to select as part of a Boolean search and asked *this one?* Participant 5, as part of a task, found *Dora*'s screenshot referred to the video she was looking for and asked *it's this one, isn't it?*

Participant 1 asked for assistance in a different manner. While trying to accomplish a task he looked at the researcher and said *see if you can do it*.

Following the interaction process, as participants watched the short video segments, some commented on the video, others sung along with the song accompanying the video. Only one issue occurred when Participant 18 successfully accomplished the type two task using the mouse and accessed the superhero cartoon *Ben 10*. He mentioned he was familiar with the video and recognised its screenshot, and as the cartoon opening scene started he said *yeah* and smiled after a few seconds, however, he then said *I'm too scared, I'm too scared* and was told to go back. He navigated directly to the red back button and clicked. After the evaluation session was over the participant mentioned one more time he got scared. It seemed as though the participant was familiar with the *Ben 10* character but had never watched the cartoon.

Other participants, on the other hand, were so familiarised with the video segments that they knew them by heart. Participant 3 accomplished the type three task with the mouse and accessed the *Ratatouille* video. The video fragment was the trailer for the *Ratatouille* film and as the video started the participant said *I have this one in my Cars DVD*, referring to the Disney film, *Cars*, in which the *Ratatouille* trailer was released. He laughed watching the video and said the lines in the trailer ahead of the characters. Participant 5 also commented on a video segment and told the researcher what was about to happen in the scene, but in this case, she remembered the details of the video she had chosen for the free choice task. In addition, while she was filling out the Fun Toolkit survey she said *I want Cebolinha again, Cebolinha*, referring to the character of the video she just watched on the free choice task, providing evidence that she enjoyed repetition.

5.3.5 Summary of the Results

The results presented are summarised in Tables 21 and 22 below, indicating: factors that were found to have influenced participants' interactions with the iTV prototype (Yes); factors that also had an influence on the interaction but whose effect was less significant (Slightly); and factors that did not seem to have influenced participants' interactions (No).

	Tasks Accomplished	Time to Accomplish Tasks	Hints Needed	Interactions Performed	Preference
Age	Yes	Slightly	Yes	Slightly	Slightly
Gender	No	No	No	No	No
Country Context	Yes	No	Yes	Slightly	Slightly
Condition	No	No	No	No	No
Input Device	Slightly	Slightly	Yes	Yes	Slightly
Task Type	Yes	Yes	Slightly	Slightly	No

Table 21. Summary of the results for the different variables examined

Media and Device Use	Tasks Accomplished	Time to Accomplish Tasks	Hints Needed	Interactions Performed	Preference
TV	Yes	No	Slightly	No	No
Computer	Slightly	No	No	No	No
Remote	Slightly	Slightly	Slightly	No	No
Mouse	Slightly	No	Slightly	Yes	Yes

Table 22. Summary of the results by media and device use

In addition to the results presented in the tables above it was also found that participants' physical and cognitive characteristics interfered with their interaction with the prototype application. The results will be further discussed in the following section.

5.4 Discussion

Based on the results presented in this chapter, the age of the participants had an effect on the number of tasks accomplished, with older participants accomplishing more tasks than younger participants. Regarding the input device used to accomplish the tasks, the average tasks participants of both ages accomplished with each device were similar, but it was found that there was slightly more variation in the number of tasks accomplished by three year old participants when using the mouse rather than the remote. While only one three year old participant failed to accomplish any tasks using the remote, five participants of the same age did not accomplish any tasks with the mouse and among those, two participants struggled to use the device to the point where they had to be told to use an alternative input device instead. With regard to the complexity of the tasks, tasks of type one and two showed a difference in the number of tasks accomplished by three and four year old participants, but it was not as significant as for the type three tasks, in which four year olds accomplished on average significantly more tasks than three year old participants. These results indicate that younger participants accomplish fewer tasks than older participants and this difference increases accordingly with the increase in task complexity. There is also some evidence that younger participants may have more difficulty managing the mouse than the remote control. The fact is that for the simplest tasks (type one), participants were required to aim at a screenshot then click when using a mouse while with the remote only one step was required - to press the OK button when the screenshot of the video to be accessed was presented in the centre of the 3D wheel. It was then easier, especially for participants with limited motor skills, to accomplish simple tasks using the remote control rather than with the mouse.

The hand-eye coordination issues that occurred during interaction (Section 5.3.4.1) could be associated with less developed motor skills of younger participants that in turn could be related to participants' physical development and use of media and devices, in view of the fact that according to the parent's questionnaire responses, older participants watch television and use the computer on average more frequently than younger participants, and a larger number of four year olds use the mouse and remote unsupervised compared with three year olds.

The time taken to accomplish tasks using both devices is on average similar for three and four year old participants but, comparing the age groups, younger participants took slightly longer and there was more within-group variation in the time they took to accomplish tasks. Considering the task complexity, for tasks of type one and three, three year olds took longer and there was more variation in the amount of time they took, for type two tasks, the time taken by three and four year olds was similar. The average time taken to accomplish tasks by three year olds increased along with the complexity of the task, but the four year olds took longer on average to accomplish type two tasks than those of type one and three. Two possible reasons for this are that older participants learnt how to perform searches using the category buttons during type two tasks and applied it for type three, or that performing a Boolean search was too complex and only skilled participants were able to do so, bringing down the average time to accomplish tasks of this type. There is some evidence in favour of the first reason in which there is a learning curve where participants retain the information and apply it to the following task, because thirteen participants accomplished at least one of the two type three tasks and most (eight participants) actually took longer to accomplish the previous type two task than the Boolean search. Among the other participants, three had not accomplished the previous task, so only two participants took longer to accomplish a type three task than a type two. While selecting a screenshot already in display for type one tasks may be straightforward, for type two tasks, the process of looking at the buttons, selecting a category and then selecting a screenshot was relatively complex and often required hints. For tasks of type three, however, participants had already understood the process and selected two categories in less time than they had taken to select only one in the previous tasks. The fact that the same does not hold true for younger participants, however, could be related to their less developed motor skills and frequency of media and device use, so that despite already knowing how the process of selecting categories worked, the effort required to interact via the input devices to select two categories, instead of only one, still resulted in them taking longer to accomplish type three tasks.

It was found that on average older participants were given more hints, but this data should not be interpreted generally as showing that older participants required more hints to accomplish tasks, because only hints given to participants during tasks that were actually accomplished were considered. Older participants accomplished more tasks so were consequently given more hints. Calculating the total amount of hints given, including those for tasks that were not accomplished, would also not elucidate how the number of hints needed could be related to participants' age, because nine participants were found to be tired or fidgety and the session was interrupted before the last task. Seven of them were three year olds who were then given no hints for the last task only because they were not asked to accomplish it. When the number of hints given to accomplish tasks was analysed in detail, a similar number was found to be given for mouse and remote navigation in the two age groups. Considering the complexity of the tasks, for tasks of type one and two the number of hints given to three year old participants was on average higher than those given to four year old participants, and for type three tasks, the average number of hints given to participants of both ages was similar. As a result, younger participants needed more hints on average than older participants. It can also be seen that more hints were given to three year old participants to accomplish the two type two tasks compared with tasks of type one and three, while four year old participants needed more hints for the type two tasks compared with those of type one. Additionally, for this age group, the number of hints needed to accomplish tasks two and three was similar but less variation was found for type three tasks. For type one tasks, the hints were mostly regarding navigation using the input devices. For tasks of types two and three, there were additional hints for participants to look at the buttons on the bottom of the screen, then hints for the categories to be selected. For type three tasks, two categories had to be selected instead of one, increasing the chances category hints would be required. There was no evidence, however, that participants needed more hints to accomplish tasks of type three rather than tasks of type two. This emphasises the finding reported previously that participants learned how the interaction worked during type two tasks and subsequently needed less time and fewer hints to accomplish the following task.

Regarding the number of interactions registered, clicks and buttons pressed, three year olds performed on average a larger number of interactions than four year olds in their accomplishment of tasks. The recording of only mouse clicks, as opposed to mouse movements, meant that the minimum number of interactions required in order to perform tasks using the mouse was significantly lower than when using the remote. Participants were only required to aim at the target elements, buttons or screenshots with the mouse, while with the remote they had to navigate to the target elements using the arrow keys, in which every button pressed was counted. As a result, four year olds needed significantly fewer interactions to accomplish tasks using the mouse than when using the remote. For younger participants, however, the same did not hold true. Three year olds performed a similar number of interactions using the mouse and the remote control. This could also be related to younger participants' limited motor skills and less frequent media and device usage. The navigation with the remote needed more interactions but there was no requirement to aim. Aiming and clicking on the target using the mouse could be strenuous, and three year olds would then need a large number of clicks to hit the target they wanted, so the number of interactions performed with both devices was even. For type one tasks, three year olds performed significantly more interactions than four year olds. For tasks of type two and three, the number of interactions performed to accomplish tasks by three and four year old participants was similar. However, it is important to highlight that for half of the three year old participants, only type one tasks were accomplished.

The data collected with the adapted Fun Toolkit survey could not be triangulated, and it is not evident whether preschool children were unable to rate their experience with a specific device and compare two devices, or whether they wanted to ‘play fair’ and balanced their responses. Nevertheless, the overall degree of preference for both devices was higher for four year olds than for three year olds and both three and four year olds’ average degree of preference for the mouse was slightly higher than for the remote.

Considering the gender of participants, a difference was observed during the video analysis between boys and girls in Brazil, in which the boys sounded more confident during the interaction than the girls. Following a series of calculations, however, it may be seen that the number of tasks of the three different types accomplished by boys and girls using the mouse and the remote was very similar. There was also no significant difference in the time boys and girls took to accomplish tasks, the number of hints they needed or the number of interactions performed. The degree of preference between the two genders for the remote and the mouse was also similar. As a result, it can be said that boys and girls may have expressed themselves differently in the evaluation sessions conducted in Brazil but, unlike in Weeramanthri’s (2008) study, the gender of the participants was not fount to interfere with the interaction process.

Regarding participants’ country context, it can be seen that participants in Brazil accomplished more tasks using both the mouse and the remote, but they were on average older than participants in the UK. Considering the task types, participants in Brazil and in the UK accomplished a similar number of type one tasks, but for tasks of type two and three, participants in Brazil accomplished on average more tasks than those in the UK. The time participants in both countries took to accomplish the tasks was similar. More hints were given to participants in Brazil but, as previously mentioned, the only hints calculated were those given to tasks which had been accomplished. Participants in Brazil accomplished more tasks, so they were given more hints accordingly. When hints were divided into different types it was found that a similar number of mouse and remote navigation hints were given to participants in both countries. For the different types of task, a slightly larger number of hints to accomplish type one tasks were given to participants in the UK compared to the number given to participants in Brazil. For tasks of type two, the number of hints given to participants in both country contexts was similar and for type three tasks, a slightly larger number of hints were given to participants in Brazil rather than participants in the UK. Participants in Brazil performed on average more interactions to accomplish tasks than those in the UK, but this was also because they accomplished more tasks. The average number of interactions (clicks) to accomplish tasks with the mouse was similar for participants in both countries. For tasks accomplished using the remote, however, participants in Brazil performed significantly more interactions (pressed more buttons) than participants in the UK. For type one tasks, participants in the UK performed on average more interactions to accomplish the tasks than participants in Brazil. For tasks of type two and three, participants in Brazil performed more interactions than participants in the UK, but again, this was mainly because participants in Brazil had accomplished more tasks on average.

Participants in Brazil may have accomplished more tasks because they were older. Another factor that could have influenced this result is that they watch more television than the participants in the UK, according to the responses from the parents' questionnaire, so they may have found it easier to accomplish tasks because they were more familiar with the videos to be found, recognised the categories to which they belonged and the screenshots representing them. Another aspect that could have had an effect on the number of tasks accomplished is that a larger number of participants in Brazil used the remote control with no supervision compared to participants in the UK. Their familiarity with the remote may have made them more comfortable using the device than participants in the UK. The computer and mouse usage for both country contexts was similar. Other aspects that could have influenced the number of tasks accomplished by participants in both country contexts were noticed during video analysis. It was observed that participants in both countries behaved differently. Participants in the UK, especially the younger ones, showed more signs of tiredness and restlessness, so the evaluation sessions had to be interrupted before they were asked to accomplish all six tasks. In this case, the interruption of the evaluation session may have been related to the participants' age rather than to their country context. Nevertheless, participants in the UK asked to watch videos of their own choice instead of the videos to be found as part of the tasks, while participants in Brazil did not question the task to be accomplished. Among several possible reasons for this is the fact that in Brazil nursery activities are more structured and directed while in the United Kingdom children are usually given the choice of several different activities and play-areas that they are able to choose according to their abilities and preferences (Bright, 2009). This is reflected in the participants' behaviour during the evaluation sessions. Children in Brazil followed the tasks as they were used to doing during activities in the nursery while children in the UK asked to choose their own videos instead.

According to the Fun Toolkit responses, it was noticed that the degree of preference for participants in Brazil and in the UK for the mouse was similar, but for participants in Brazil the degree of preference for the remote appears to be higher than for those in the UK. For participants in Brazil and in the UK alike, the degree of preference was slightly higher for the mouse than for the remote.

There was no difference found on the number of tasks accomplished, hints given to accomplish tasks, interactions and participants' preferences for the four different conditions to which participants were submitted. It was noticed, however, that there was a difference in the time taken to accomplish tasks with each device: participants on conditions one and three, who had used the mouse first and remote second, took longer to accomplish tasks with the mouse than with the remote; while participants on conditions two and four, who had used the remote first and the mouse second, took longer to accomplish tasks with the remote than with the mouse. This is another demonstration of the learning curve that occurred during the evaluation sessions and the effect it had on participants' interactions with the prototype.

The interactions with the two input devices, remote control and mouse, were analysed in detail, and it was found that on average the same amount of tasks were accomplished using the remote as when using the mouse. There was, however, slightly more variation in the number of tasks participants accomplished with the mouse compared to the remote. As previously discussed, it was found that it was easier, especially for younger participants, to accomplish simple tasks using the remote than using the mouse. With regard to the complexity of the tasks, on average more tasks of type one and two were accomplished using the remote control than using the mouse, while for type three tasks, which involved the selection of two categories; more tasks were accomplished on average using the mouse than using the remote control. This provides evidence that it was actually easier to perform simple tasks with the remote compared to with the mouse. When the task complexity increases, however, it becomes easier to accomplish tasks using the mouse because the navigation with the remote requires too many steps, navigating to two different buttons using the arrow keys then selecting it by pressing the OK button, in comparison to the comparatively simpler process of aiming and clicking with the mouse. For this reason as well, participants accomplished tasks on average slightly faster using the mouse than they did using the remote. On average a larger number of hints were given to participants to accomplish tasks using the remote than to those using the mouse and the difference increased with the increase in task complexity. Navigation with the mouse seemed more intuitive; although some participants did ask if they should click on screenshots for the first task, for the following tasks most did not need any hints and understood the way the navigation worked and clicked on buttons and screenshots without hints. For the remote, on the other hand, there was not a significant difference in the average number of navigation hints for each task type. Participants needed hints on how to navigate with the remote throughout the three tasks, so it seemed that, differently from the mouse navigation, they did not memorise as quickly how the interaction with the remote worked. The average number of interactions performed to accomplish tasks using the mouse was significantly lower and presented significantly less variation than the average number of interactions performed to accomplish tasks using the remote. As discussed previously, the mouse movements in the process of aiming at targets were not recorded, only the mouse clicks, while all buttons pressed to navigate using the remote were summed. As a result, the number of interactions with the remote was significantly higher. For type one tasks, however, the process was similar for both devices; only one click and one button press were required. Participants could either wait for the screenshot of the video they wanted to access to appear on the centre of the wheel and aim and click with the mouse or press the OK button using the remote. Only one interaction would be needed no matter which device were used. Some participants did as above, but most participants did not wait for the screenshot to appear but actually navigated through the screenshots using the arrow buttons and then pressed the OK button. So even for the simplest tasks, the number of interactions performed was on average higher using the remote than using the mouse as an input device.

The two alternative input devices to the mouse, the touch pad and keyboard, were used by only three participants. One participant in the UK chose to use the touch pad instead of the mouse. In Brazil, a participant also asked to use the touch pad, but unfortunately he was excluded from the study because the computer screen was positioned incorrectly so that the webcam only captured his forehead. However, the fact that he asked to use an alternative input device presents some evidence that there are children who do prefer to use the touch pad instead of the mouse because they are more accustomed to it. He mentioned, as did the participant in the UK, that his parents’ computer had a touch pad instead of a mouse. The participant who used the touch pad accomplished the three tasks with the device, took on average slightly more time than those who used the mouse, but needed only one touch pad navigation hint and fewer interactions than participants who used the mouse. For the participant who used the touch pad, the interaction was then very similar to that of participants who used the mouse. There was a different situation, however, for the two participants who use the keyboard. Firstly, they did not choose to use it, but were told to do so because they struggled to interact with the mouse and touch pad. It is likely that they struggled to interact because of hand-eye coordination issues, mostly related to their developmental level, since they were the younger participants in the study; one was three years and one month old, the other three years and two months old. They accomplished only one task, the simplest, took significantly longer to accomplish the task than participants who used the mouse, were given on average two keyboard navigation hints and performed significantly more interactions to accomplish the task than participants who used the mouse. The keyboard made the task accomplishable for these participants, but certainly did not make it easy. They struggled to use the keyboard arrow keys to navigate through the screenshots and pressed the keys repeatedly and found it hard to stop at the screenshot they wanted. However, when they managed to do so, they pressed the space bar and accessed the video. The fact that the big space bar button gives access to the video facilitated the interaction and even younger participants, excluded from the study because they were younger than three years old at the time of the evaluation session, were able to interact. They did not use the arrow keys at all, but simply waited for the screenshot of the video they wanted to access to appear on the middle of the wheel and pressed the space bar. If they wanted to return to the menu they pressed the space bar one more time. They seemed to be able to interact with the prototype application using only one button.

The different task types were also analysed in detail and it was found that for the type one tasks, where participants were asked to select a video already in display, the accomplishment rate was significantly higher than that for type two and three tasks. Nevertheless, for type two tasks, where participants were asked to perform a simple search selecting one category, the accomplishment rate appears to be similar to that for type three tasks, where they had to select two categories. It can thus be seen that increasing the complexity of the task from selecting one category to selecting two categories to find a video did not result in a reduction on the number of tasks accomplished. This is another indication that as a participant learns how the interaction process works, s/he is able to perform both simple and Boolean searches with the same facility. The type one tasks were accomplished significantly faster than tasks of type two and three. The average time participants took to accomplish type two and three tasks was similar, but there was more variation in the average time to accomplish type two tasks than there was with type three tasks. On average, a smaller number of hints were given to participants to accomplish type one tasks when compared to the number of hints given to participants to accomplish tasks of type two and three. The average number of interactions performed to accomplish type one tasks was lower and presented less variation than those for tasks of type two and three. Participants were able to accomplish the simplest tasks more quickly, with fewer hints and interactions. Despite the level of complexity increasing from type two tasks to those of type three, however, there was no significant difference in time, hints or interactions necessary to accomplish the two different types of task. These results show once more that learning occurred during the evaluation sessions and participants were able to accomplish Boolean searches in a similar time, with a similar amount of hints and number of interactions as when they accomplished simple searches.

Analysing the data gathered through the parents' questionnaires, certain findings can be discussed. It was noticed that participants who watched television more frequently, on a daily basis, accomplished more tasks than participants who watched TV less frequently, on a weekly basis. However, the time they took to accomplish tasks, the number of interactions performed and the degree of preference for the remote were similar. It was found that participants who watched television less often were given on average fewer hints than participants who watched television more often. It should be highlighted, however, that in all likelihood, this happened because, as in situations described earlier, more hints were computed to participants who accomplished more tasks, in this case the ones who watch television more often. With regard to computer usage, it was also noticed that participants who used the computer on a daily basis accomplished more tasks than participants who used the computer only occasionally. In this case, the time taken to accomplish tasks, hints needed, interactions performed and degree of preference for the mouse were similar.

Data from participants' device usage suggest that participants who have unsupervised use of the remote control accomplished on average more tasks, accomplished tasks more quickly and needed fewer remote navigation hints than those who did not use the device or used the remote only with supervision. There was, however, no statistically significant difference in the number of interactions performed to accomplish tasks and in the degree of preference for the remote control between participants who used the remote unsupervised, those who did not use it at all and participants who only used the device with supervision. Participants who had unsupervised use of the mouse accomplished on average more tasks, were given fewer mouse navigation hints, performed significantly fewer interactions and had a higher degree of preference for the mouse than participants who did not use the mouse, yet the time they took to accomplish tasks was similar.

It can thus be said that participants who watched television and/or used the computer more frequently accomplished more tasks. With regard to the input devices, participants who used the remote control and/or mouse unsupervised accomplished on average more tasks and needed fewer navigation hints than those who did not use the device or used it only with supervision.

Through the Fun Toolkit survey it could be noticed that the degree of preference for the remote was on average lower for participants who did not accomplish any task with the device. The same occurred for the mouse, with the degree of preference for the mouse being slightly lower for those participants who did not accomplish any tasks using it. However, evidence was not sufficient to show that the degree of preference for each device was different according to the time taken to accomplish the tasks with the devices, the number of hints given to interact with each device or the number of interactions performed to accomplish the tasks.

The results suggest that participants who used media more frequently and devices independently were able to perform more tasks and because overall they had a better experience interacting with the prototype using the input devices provided, their ratings for the devices were higher.

It was already stressed that screenshots originating from promotional material or from video segments were recognised by participants as representing videos. Some participants failed to recognise a few screenshots (Section 5.3.4.2.1), but the fact that they selected another screenshot from the same category shows that the screenshot carries enough information to indicate to which category its video belongs.

It was noted that the icons representing categories that were part of the tasks to be accomplished were recognised by at least one participant (Section 5.3.4.2.2). Some icons were recognised by more participants than others, but participants were able to roll over the buttons to identify the categories so that the icon recognition was not as crucial for the interaction process. There were two issues that occurred during the interaction – both of which can be easily resolved. The icon for the category Nick Jr. was an elephant and confounded with the ‘animals and nature’ icon. The channel does have other logo options that could be used instead. The films *Cinderella* and *Ratatouille* could be included in the Disney category along with all Disney productions, instead of just programmes from the channel Playhouse Disney, since participants related the films to the Disney brand.

For the coloured buttons, it was noted that the red ‘back’ button represented by an arrow was recognised by some participants; others needed hints to use it for the first time, but subsequently learned to recognise it and were able to use it in the following tasks. There was an issue in Brazil with the red ‘close’ button. It was interpreted as a programme logo instead of ‘close’ so, for that audience, it would probably be interesting to review if another icon should be chosen or the platform standards maintained (RL27). In the UK, the button was recognised and used for its proper function.

There is not enough evidence that participants understood the ‘favourites’ feature of the prototype. Only one participant actually used it and then verified the programme was displayed on the ‘favourites’ button. There is certainly a need for the feature though, because during the evaluation sessions participants talked about the programmes they liked and they gave the impression they liked repetition. Therefore, adding videos to ‘favourites’ would be a way in which they could select the videos they liked and facilitate repeated access thereto. To test the appropriateness of this function’s presentation would, however, require a longer evaluation session or a specific session to test solely this prototype feature.

It was also found that there is a need for the ‘clear all’ button. Some participants turned on too many buttons and having another button to turn every button off at the same time would certainly seem to be useful and time saving. However, there was no evidence that participants understood the function of the button. This feature would also have to be tested in an additional session.

The help button was recognised and used, but it was not clear which type of assistance participants were looking for when they accessed it. The help segment was misinterpreted by some participants who tried to interact with it, so it would probably be useful to have a character, for instance, explaining the interaction process in such a way that it would not be as boring to watch and that would clarify that the buttons displayed on the screen were inactive because they were part of the help segment and not the prototype itself.

The hints given are related with the assistance needed that could then be provided on a help section. Among the five different types of hint, the number of remote control navigation hints was significantly higher than the other types concerning mouse navigation or categories. This shows that users would benefit from further assistance on how to navigate using the remote. It would almost certainly be worthwhile illustrating the device in the instructions because it could be that participants did not relate the arrow keys or the OK button mentioned during the tutorial with the actual buttons on the remote. If the buttons were actually shown, it is likely that fewer hints would be needed.

It is important to highlight that during the interaction with the prototype, there was no evidence that participants understood they were performing a Boolean search for type three tasks. Participants did not demonstrate comprehension that by selecting two buttons they were performing a conjunctive search and narrowing down their search result. Additionally, whereas there were participants who did not require any hints to accomplish tasks of type one and two, all participants needed at least one hint to perform type three tasks. With this in mind, it could be interesting to emphasise the process of performing a Boolean search on the tutorial/help segment.

No participants were able to recognise the pause icon. The reason for this could be that participants were not familiar with the pause icon itself or that they just did not recognise it because it was white and displayed at the bottom left corner of the screen. The button could be placed at the right hand corner of the screen, and the icon displayed in a coloured box to make it more evident, but it would have to be investigated whether children actually understood the symbol used and also if they actually desired an option to pause videos.

Participants experienced problems differentiating when buttons were on and off, and so to ameliorate these visual issues, it could be that the change between states has to be less subtle than a simple lighting up of the button, but could extend as far as changing the colour of the entire button for its different states.

Some participants tried to interact with the greyed inactive buttons and it may be interesting to have a form of audio feedback that lets users know that the button is inactive and cannot be turned on.

Two participants made comments when the prototype was opened at the beginning of the evaluation session. One used the expression “wow” that suggest she was impressed with the prototype application, and the other participant said she had a computer like that. This participant, previously on the session, mentioned she has a toy laptop, so this suggests that the prototype actually looked for the participant like a toy laptop and is then tuned to children’s universe.

Overall, participants understood how the interaction with the prototype worked and provided evidence of this by interacting with no or very few hints, and talking about the interaction process. It was found that the wheel movement could be a problem. It was great to have the wheel moving with the screenshots, especially to provide interaction via one click/button press only, where participants could wait for the screenshot of the video they wanted to access and click or press the OK button on the remote. Nevertheless, when the cursor was moved away from the wheel, it sped up and sometimes caused problems. Ideally, the movements of the wheel would be varied as is, but the wheel would move slower as the cursor approaches it and would gain in speed when the cursor moves away from it, but only towards an established speed limit that would not compromise the children’s ability to select the desired screenshot.

Following the interaction process, only one issue occurred as a participant was watching a video segment. Although he mentioned he was familiar with the *Ben 10* superhero and recognised its screenshot, after a few seconds of the video he said he was too scared. The video was selected because it was found to be very popular among preschoolers during the previous stages of the research, but it was certainly disturbing to this participant. An interesting point is that as he was told to go back, he navigated directly to the red back button and clicked. This shows that children are able to control the prototype on their own and would not only be capable of selecting the videos they would like to watch but also of quickly stopping the viewing of video content that disturbed them in any way. They would have control of their viewing experience.

5.5 Conclusions

It may be concluded from the findings presented and discussed that age influences user experience, having an effect on the number of tasks accomplished, the time taken to accomplish tasks and the number of hints and interactions needed. Among three and four year old children, younger participants were able to accomplish fewer tasks than older participants, especially complex tasks. For younger and less experienced participants, the remote proved to be more effective than the mouse, particularly to accomplish simple tasks. Younger participants took slightly longer to accomplish tasks compared to older participants. In addition, younger participants performed more interactions to accomplish tasks than older participants. They performed a similar number of interactions using the mouse and the remote while older participants performed significantly fewer interactions to accomplish tasks with the mouse compared to the remote. These differences in the way three and four year old users interact may be associated with the less developed motor skills of younger participants that could in turn be related to the participants' physical development combined with their less frequent use of media and devices. Overall, it is believed that older participants had a better experience interacting with the prototype using the remote control and the mouse and this was reflected in their higher degree of preference for both devices compared to younger participants.

Gender was not found to have an effect on user experience. The way boys and girls interacted with the prototype was similar; they accomplished on average the same amount of tasks and took a similar amount of time, hints and interactions to accomplish the tasks. The only difference noticed was that within the group in Brazil, boys sounded more confident than girls, but this could simply be attributed to a different manner they used to express themselves during evaluation sessions.

The user experience was different for the two country contexts; participants in Brazil seemed to have a better experience interacting with the prototype and accomplished more tasks than the participants in the UK. Nevertheless, this difference could actually have been influenced not by the country context itself, but by age, frequency of media and device use or type of curriculum followed by the nurseries in each country.

The conditions participants were submitted to had an effect only on the time taken to accomplish tasks, with participants taking longer to accomplish tasks using the first interactive device they were given than using the second device. This is one piece of evidence confirming that participants learnt how to use the prototype and applied their acquired knowledge to the following tasks, enhancing their experience throughout the session.

It was found that participants had a better experience performing simple tasks with the remote compared with the mouse. As the task complexity increased, however, participants had a better experience using the mouse, a device with which they were able to more quickly accomplish tasks, requiring fewer hints and significantly fewer interactions than they would have needed had they been using the remote. Additional input devices were useful in catering for participants who preferred to use another device or were not able to use the mouse. Children who are used to the touch pad may prefer to interact with it instead of the mouse and children who are unable to use the mouse may interact with the keyboard with which no aiming is necessary.

The interesting fact noticed about the task types was that participants accomplished the simplest tasks (type one) faster and needed fewer hints and interactions, but there was no significant difference in the time taken to accomplish tasks, and the number of hints or interactions for tasks of type two and three, although the level of complexity increased from type two to type three tasks. This is additional evidence of the learning that took place during the evaluation sessions, enabling participants to accomplish both simple and Boolean searches in similar times, with a similar amount of hints and number of interactions.

It was concluded that participants' media and device usage influenced their experience with the prototype. Participants who watched television and/or used the computer more frequently accomplished more tasks and had a better experience than those who used the media occasionally. Regarding the input devices, participants who used the remote control and/or mouse unsupervised had a better experience, accomplished on average more tasks and needed fewer navigation hints than those who did not use the devices or used them only with supervision.

Although the Fun Toolkit results could not be triangulated, it can be seen that participants' degree of preference for each device was somehow related to their experience with that device. In general, if participants were not able to accomplish any task with a device, or they had trouble interacting with that device, this affected their experience and was reflected in their lower overall rating for that specific device. Therefore, the method could be used with preschoolers to rate their experiences after the interaction, but data should be gathered in different forms (Smileyometer, Again-Again Table and Ranking System). Results could be calculated so that every form had a similar weight and then these results could be compared with video analysis and other indicators of the overall user experience.

Some prototype elements were well understood, while others could probably be reviewed and redefined. Participants recognised screenshots and gave sufficient evidence that they understood that they represented video segments. Most icons that represented categories and functions were identified by at least one participant, but some may need to be redefined. The ‘favourite’ and ‘clear all’ buttons were found to be relevant for the prototype, but additional evaluation sessions would be required to confirm the appropriateness of their icons and functions for children. The tutorial/help segment could probably be demonstrated in a way that differentiates it from the actual prototype, and in addition, the buttons on the remote control could be illustrated and emphasis could be put on the explanation of Boolean searches. The pause button was not recognised, so further investigation is required into whether children of this age understand the icon and/or would like to use the feature. The visual feedback should be emphasised and audio feedback added to inactive buttons. The wheel movement presented some issues that could be solved by establishing a speed limit for the wheel such that screenshots could be easily chosen and selected.

All participants accomplished at least one task, proving that they were able to interact with the prototype and control it on their own using different input devices. In addition, participants were found to enjoy the interaction, especially while performing the free choice task. Hence, the prototype application meets the needs, capabilities and interests of preschoolers, and this was certainly achieved because children contributed actively towards its design.

Chapter 6. Discussion

6.1 Introduction

In this chapter the findings presented in the previous chapters are discussed and illustrated in a framework that may elucidate factors that influence preschoolers' interactions with iTV applications. The implications of such interactions are then examined, combined with the initial principles established with requirements and design decisions during the design process (Chapter 4) and the interaction issues found during the evaluation process (Chapter 5) to generate design principles for interactive television for young children. This is followed by a review of the techniques for design and evaluation used during the studies. The chapter concludes in Section 6.5 with a summary of the main research findings.

6.2 Young Children's Interactions with iTV

Age is a factor found to influence children's interactions with the iTV application during evaluation sessions. Older participants were able to accomplish more tasks, especially complex tasks, accomplish tasks faster, needed fewer hints and performed a smaller number of interactions (Chapter 5). There was, however, no indication from the closed card sorting activity that participants' age had an effect on their categorization ability (Chapter 4). These results suggest that children's physical development, that generally matures with age (Shaffer, 1999, p.165), could have affected participants' experience during the evaluation session in which they had to perform tasks using the mouse and remote control that demanded fine motor skills, but had no influence on card sorting tasks in which they had to post cards through slots in shoeboxes, a relatively much simpler task that did not require the same level of fine motor skills. Nevertheless, children's age was not the single participants' characteristic that influenced their interaction with the prototype, other factors such as media and device use and country context were also found to have an impact on the way participants interacted.

During the evaluation sessions it was found that the frequency of media use influenced the number of tasks accomplished. This could have occurred because participants who use media more frequently could be more familiar with the videos to be found, recognised the categories they belonged to and the screenshots that represented them. This hypothesis was also raised during the card sorting activity. At that stage of the research the data on media usage was not gathered systematically therefore could not be used as reference to test the hypothesis. Even so participants' ability to categorize video content according to adults' categorization standards was not found to be related to age so it could be associated with the frequency of media use: the more frequently participants use media the more familiar they are with the screenshots and the more similar the categories they form become to adults'. Device usage has also had an impact on the way participants' interacted with the prototype application during evaluation sessions. Participants who use the devices unsupervised were more familiar with the way the interaction through the devices worked, were more trained to use the skills required, as a result accomplished more tasks and needed fewer hints than those who do not use the devices or use them only with supervision. In addition, device usage is also directly related to participants' motor skills, the use of the computer mouse, for instance, is considered to promote development in young children (Meggitt, 2007, p.86). Thus, children who use input devices independently have more practice using the skills required to interact with the devices and furthermore may have such fine motor skills more developed than those who do not use input devices, or do not use them on their own. Children's motor skills are then associated with children's age and determined not only by their physical development but also by their media and device usage.

In this research participants' media and device usage was related to children's age and country context. Other studies have also indicated a similar age and media use relationship, in which older children use media more often than younger children. According to the Kaiser Family Foundation research conducted in the United States, “children between the ages of 4 and 6 years engage in most screen activities more often than those ages 3 years and under” (Rideout and Hamel, 2006, p.8). Regarding the media use and country context, it was found through the parents’ questionnaires that there was a difference in the frequency participants in Brazil and in the UK watch television but their frequency of computer use was very similar. The sample in this study is too small to elucidate tendencies of media usage across nationalities, but it may be stated that children’s use of media in different countries are likely to vary. European research had shown that despite the fact that internet use is diffused across Europe, there are still cross-national differences; in countries with overall higher internet use children are much younger when they start to use the internet (Livingstone and Haddon, 2009).

The knowledge children have about media and devices had an effect on the way they interacted with the prototype application during the evaluation sessions. As already discussed, the more frequent the use of media and devices the more familiar children may be with screenshots, categories and the way the interface and the devices work, influencing accordingly their interaction with the iTV application. This knowledge may be attained previously by using diverse media and devices or may be acquired during the interaction with the specific iTV application improving the performance on following interactions. Evidence presented in Chapter 5 confirms that participants’ experience was enhanced over the session, an indication that they may have learnt how to use the prototype application and applied their acquired knowledge to the subsequent tasks. The fact that the children’s interactions develop with knowledge and/or practice is also supported by research on use of input devices suggesting that young users experience differences in ease of use for different input devices, but performance usually improves with experience for all devices (Revelle and Strommen, 1990).

During the evaluation sessions, there were issues with children’s interactions with the input devices that interfered with their overall experience with the iTV prototype. Some participants had difficulty holding the remote control in their hands, pointing the device to the IR receiver and pressing the buttons all at the same time. Revelle

(2003) indicated that for young children the remote control is too big and it is difficult for them to keep it oriented correctly when using it from a hand-held position, therefore the author suggested placing the remote on a surface (RL42). In the research reported in this thesis such requirement was followed, but as participants rested the remote control on the plastic placed over the computer keyboard they sometimes pushed the device down and stroked the touch pad and keys interfering with the interaction. The surface provided was useful but far from ideal. In order to enhance the interaction the device would probably have to be smaller and the surface more stable and rigid so that as participants rest the remote over it there is no interference with the interaction in any way. The other issue with the interaction through the remote control was found to be the comprehension and remembrance of how the navigation works. Participants needed a larger number of remote control navigation hints compared to the other types of hints, indicating that the navigation with the remote was not as clear as the navigation with the mouse. Revelle (2003) suggests attaching shape stickers to each one of the buttons, this could be an option to make the way the navigation works more comprehensible and memorable. An alternative to facilitate the understanding and possible the recall feature of the arrow keys without adapting the device itself could be to illustrate the remote and its buttons during the tutorial/help segment. Whereas the navigation through the remote arrow keys was not so straight forward the coloured buttons were found to work well as short cuts (RL2), especially the red ‘back’ button was recognised and used by most participants during the evaluation sessions.

According to Revelle and Strommen’s study (1990) the mouse and trackball are more effective and easier for young children to use compared to the joystick or arrow keys. The mouse was proven to be the most appropriate pointing input device for children when compared with joysticks and keyboards (Jones, 1991, King and Alloway, 1992). In this study, we found an indication that the arrow keys are actually more complicated for children to interact with than the mouse. Participants overall accomplished tasks slightly faster using the mouse, needed less assistance and smaller number of interactions compared with the remote control. Nevertheless, for simple tasks and less experienced users, the remote was proven to be more effective than the mouse. The fact is, one button press interaction is ideal for young children, as research had shown (Revelle, 2003), and it may only be achieved through input devices such as the remote control or keyboard with which no aiming is necessary. To aim targets on the screen is a laborious process for young users (Donker and Reitsma, 2006), guidelines (RL18 and RL21) were followed during the process of design of the prototype application in order to make the task easier. The buttons were placed close to each other on the screen, but distanced enough to compensate for inaccuracy in targeting (Chiasson and Gutwin, 2005) and each button had a diameter of at least 64 pixels (Hourcade et al., 2004). Such guidelines may have improved children's interactions with the interface using the mouse, there were, however, some participants who still faced difficulties aiming the targets and others that were not able to do so. They were, on the other hand, capable of using the arrow keys and OK button on the remote or arrow keys and space button on the keyboard providing evidence that for young and less experienced users to accomplish simple tasks, the navigation with arrow keys can be more effective and easier compared to the mouse. Moreover, touch screens could be even more efficient than the remote control for youngest users because it would remove difficulties on pointing and operating devices, research suggests that "touch screens are always well received by young children" (Hourcade, 2008).

It is worth highlighting that there was no evidence during the evaluation sessions suggesting that the mouse size was an issue that affected the interaction. This result is in line with previous research suggesting that mouse size does not affect performance (Hourcade et al., 2007). The device used in the present study was smaller than a regular adult sized mouse but slightly bigger than a notebook mouse and, in contrast with the remote control; participants did not present any problem that could be directly related to the size of the mouse.

Participants' age, country context, motor skills, media and device usage influenced during the evaluation sessions the number of tasks accomplished, the time taken to accomplish tasks, complexity of tasks accomplished, assistance needed and interactions performed to accomplish tasks. The combination of such factors determined the overall user experience using the interface that influenced the degree of preference for the input device used to interact.

As described in Chapter 5, the time participants took to accomplish tasks was measured in seconds; the assistance based on the number of hints required and the interactions was the sum of clicks or buttons pressed. Regarding the complexity of the tasks, there were three types of tasks: for tasks of the type one participants had to select the video already on display, for tasks of the type two they had to perform a simple search selecting a button to find the video and for tasks of the type three they were required to perform a conjunctive Boolean search selecting two buttons to find the video. It was observed that participants were able to perform all types of tasks, inclusive Boolean searches, but there was not enough evidence they comprehended they were creating conjunctive Boolean searches during the process. Hutchinson (2005) analysed children's comprehension of Boolean searches on simultaneous and sequential interfaces and concluded that older children (10 to 11 years of age) understood they were creating a conjunctive Boolean query significantly more than younger children (6 to 9 years old). In addition, she stated that "younger children require more time and more hints than older children to find two categories, regardless of the interface" (Hutchinson, 2005, p. 160). Thus, it is not evident if during the evaluation sessions the Boolean queries were understood by such young participants, but if children are given time and assistance they can actually accomplish complex tasks and perform Boolean searches that may narrow down the list of videos to choose from, resulting on a better user experience while searching for videos to watch.

6.2.1 Framework of Children's Interactions

According to the findings presented above (Section 6.2), participants' interactions with the prototype application during the evaluation sessions suggest factors that influence young children's interactions with iTV through input devices. These factors are illustrated in the framework below.

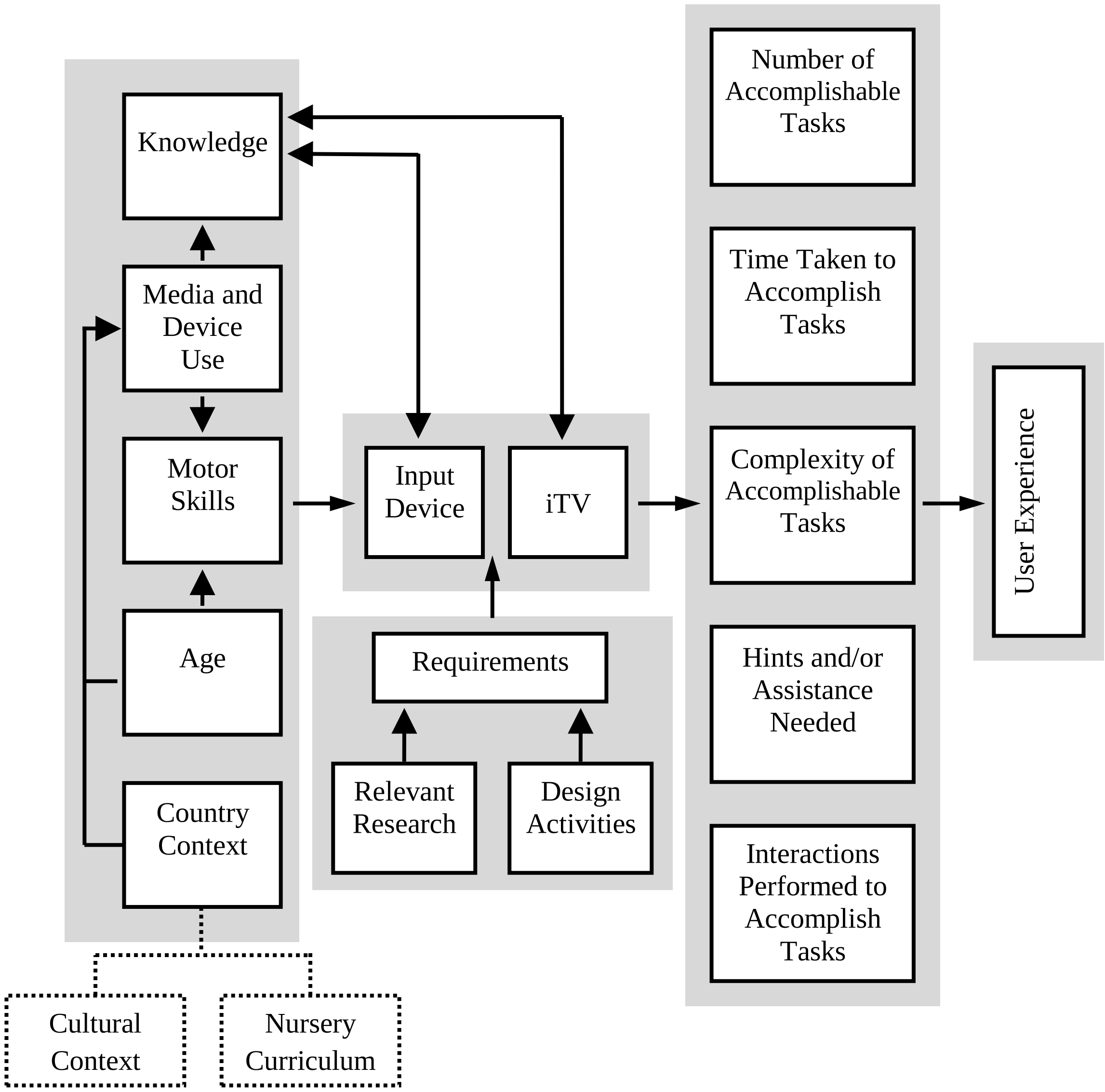


Figure 84. Framework of children's interactions with iTV

Legend:

Young children's interactions with iTV through input devices are influenced by: children's age, motor skills, country context, media and device use and knowledge.

Motor skills are affected by children's physical development associated with their age and their use of media and devices.

Children’s media and device usage varies according to children’s age and country context: children of different ages (in this case three or four years) in different countries (in this study Brazil and United Kingdom) have different patterns of media and device usage that in turn influence the knowledge they have about the media and the devices. Indications were found that suggest cultural context and nursery curriculum vary according to the country context. Country context incorporates several other variables, such as environmental characteristics, beyond the scope of this study, but these two factors were found to have an influence on the interpretation of icons and how participants behave during evaluation sessions. Knowledge, of how both the application and the input device work, is also gained as children interact with the iTV application through the input device, resulting on an improvement on their performance during subsequent interaction. These factors influence the way the interaction with iTV occurs along with the input device used for interaction and the characteristics of the iTV application. Those should be defined by requirements gathered and design decisions made based on the existing research and design activities. As a result, the user characteristics, input device and iTV application features determine the number of accomplishable tasks, time taken to accomplish tasks, the complexity of the tasks that can be accomplished, the hints and assistance as well as interactions needed to accomplish the tasks. The combination of such aspects is reflected in the user experience.

6.3 Design Principles

Design principles are defined by Jan van den Akker (1999) as heuristic statements of a format in which “if you want to design intervention X [for the purpose/function Y in context Z], then you are best advised to give that intervention the characteristics A, B, and C [substantive emphasis], and to do that via procedures K, L, and M [procedural emphasis], because of arguments P, Q, and R”.

In accordance with this definition, in order to design iTV interfaces for young children to interact without adult help it is best to give the interface some characteristics that facilitate interaction for this age group and to do that it is important to develop the interface based on requirements and design decisions supported by relevant research and design activities, because it is essential that the interface meets children's' needs and their needs may be understood by reviewing the literature or involving them in the process of design, ideally both.

In the work described in this thesis the design principles evolved from the process of design and evaluation of the prototype application. The requirements based on existing relevant research and design activities informed the prototype and the features the input devices should have had. Following the evaluation of the prototype the list of requirements[6] was reviewed and complemented with empirical evidence to be validated into design principles. These design principles are divided into principles for input devices (Section 6.3.1) and principles for iTV application interfaces (Section 6.3.2) and presented below. The principles are not intended to be a recipe for success but to offer assistance to other researchers to develop the most appropriate iTV interfaces for three and four years old children.

6.3.1 Design Principles for Input Devices

In this section the requirements for the input devices are reviewed and, supported by the results from the evaluation sessions, evolve into principles to be used when choosing or designing input devices for young children to interact with iTV applications.

The mouse and remote control were used as input devices for interaction with the prototype application during the evaluation sessions. It was noticed that for the mouse to be easy to use by young children it should, if possible, have only one

6 The acronyms used for the requirements were: Requirements from literature (RL), Requirements from existing applications (REA), Requirements from observing children (RO), Requirements from card sorting activities (RCS), Requirements from low-tech prototyping sessions (RLTP), Requirements from experts' evaluations (RE), Requirements from prototype adjustment session (RPA).

button because children did not encounter any problems to click using the one button mouse, for regular mice to mimic a one button mouse both buttons should have the same functionality (RL12).

P1 – Use a one-button mouse as input device

To click appeared to be straightforward but to aim at targets was a more complex process. There were participants who faced difficulties in aiming the mouse at the buttons and screenshots on the screen, others who were not able to do so. For this reason it is important to enable interaction via alternative devices such as the keyboard (RO4) and touch pad (RPA1) to accommodate children's varying abilities and fine motor skills.

P2 – Enable interaction via alternative input devices

Regarding the remote control it was noticed that a limited number of keys should be used (RL13) so that the range of finger movements required is limited. In addition, it was found that along the evaluation session participants were able to memorize the buttons used for interaction, if too many buttons were used it would be harder for children to memorize the function of each one.

P3 – Use a limited number of keys for interaction with the remote control

The use of colour buttons for interaction should be explored on the remote control (RL2), in view of the fact that during the evaluation sessions it was proven to be an effective way to interact. Some participants recognised the back red button straight away and related it to the red button on the remote; others did need hints to use it for the first times, but learnt, recognised and used it during following tasks.

P4 – Explore the use of coloured buttons

During evaluation sessions, participants needed a significantly higher number of remote control navigation hints compared to the other types of hints that suggests that users could benefit from assistance on how to navigate using the remote. The interaction via colour buttons appeared to be clearer than via the arrow keys and OK button (RL14), most likely because the colour buttons were presented on both screen and remote control while the arrow keys were only presented on the remote. For this reason it is suggested to present the arrow keys illustrated on the screen, at least during the tutorial/help segments, so that children can relate it to the buttons to be pressed and comprehend how the interaction works.

P5 – Illustrate all buttons used for interaction on the screen

For very young users, early threes, it is useful to enable one button press interaction in which they can use a single button to interact with the application. The fact that the OK button on the remote control and space bar on the keyboard gave participants access to the videos they wanted to watch and the possibility to go back to the menu offered them a good user experience without demanding complex interactions that would require fine motor skills.

P6 – Enable one button press interaction

While the mouse size was found to have no influence on the interaction participants had problems to hold the remote control and needed to rest it on the plastic placed over the computer keyboard. This indicates that a smaller device would probably be best for children to interact (RL1).

P7 – The remote control should be small enough for children to hold comfortably

Some participants who struggled to interact with the input devices tried to touch the buttons and screenshots they wanted to access by touching the computer screen. This could indicate that touch screens or tangible technologies in which children would be able to direct manipulate screen objects could also be a way to offer very young users a good experience. But such technologies were not used during evaluation sessions, so there are no empirical arguments in their favour, they would have to be tested to become design principles.

6.3.2 Design Principles for iTV Interfaces

In this section the requirements used to develop the prototype application are reviewed and supported by the results from the evaluation sessions to suggest characteristics iTV interfaces should have to provide users with a good experience during interaction.

Flexibility was found to be an important feature of the system (RL3). The wheel movement, for instance, allowed interaction via one click/button press only, and the possibility to vary the movement speeding up or slowing down the wheel catered to both expert and novice users. Nevertheless, it was found important to establish a limit in which the wheel never goes too fast so that in any speed children would still be able to select the screenshot they want. Thus, the system should be flexible and efficient (RL29), tailored to the entire range of target users.

P8 – The system should be flexible and efficient; provide alternative ways of interaction to cater to the entire range of target users

The interface was composed by an illustration of a device, inspired by technology for children, in which the buttons looked clickable (RL24). Participants seemed to comprehend instantly the interactive elements of the interface therefore such features should be incorporated into interfaces for this age group.

P9 – Interfaces should be inspired by things children are familiar with; make clickable items look clickable

On each button categories were represented by icons that were metaphors associated with simple words (RL23). However no text was used (RL6), the categories' and icons' labels were presented by audio feedback. During evaluation sessions, some icons were recognised by more participants than others, but the function of the buttons could be identified through the audio feedback. For this reason, the interaction was facilitated when the icon was recognised but if the icon was not recognised it did not present major problems for the interaction process.

P10 – Use metaphors associated with labels to represent icons
P11 – Use audio feedback instead of text

It was found through the literature that icons should have at least 64 pixels diameter (RL21) and should be placed close to each other but distanced enough to compensate inaccuracy in targeting (RL18). In addition, in order to enhance target acquisition the cursor activation area should be enlarged (RL22). Such activation area should be as big as possible without generating confusion by interfering with other targets or being too distanced from the original target impeding the association. During the evaluation sessions, most children were able to interact with the prototype that took into account such requirements. There were participants who faced difficulty aiming at targets on the screen, but those would probably need extremely larger targets placed significantly far from each other that would imply in presenting very few items on the interface at each time. For this reason, it is suggested to follow the requirements above (RL21, RL18 and RL22) that will provide a usable interface for most users, and add flexibility to the system (P8) to cater to those who might, for instance, find difficult to aim at targets using the mouse.

P12 – Make icons with at least 64 pixels diameter, place them relatively close to each other and make the cursor activation area as big as possible without causing confusion

The important icons should be placed in the middle of the page (RL26). The cursor should move logically, and its initial position should be close to important links (RL20). It was found that by placing the screenshots that gave access to video segments in the middle of the page and positioning the cursor over the screenshot in the centre of the wheel the one-button interaction was made possible and the interaction overall was facilitated as participants located the cursor immediately.

P13 – Place important icons in the middle of the page and make the initial position of the cursor close to them

In order to situate the user clarifying the navigation, on-screen selectors should be prominent (RL19), audible and visible feedback should be provided through animation and audio rollovers indicating functionality (RL35). It was already mentioned that the audio feedback played an important role in the interaction process and should be used on interfaces for preschoolers (P11). It is important to note that to avoid interference there should be a 0.5 second delay on the audio feedback (RE6). In addition to the audio feedback the prototype incorporated a visual feedback. An animation indicated when a button was selected; the button would glow when turned on and would go back to its original stage when turned off. It was found, however, that some participants did not differentiate the three instances of the buttons, for this reason the visual feedback should not be as subtle as lighting the button up. It is important that the visual feedback indicating different states of buttons and links is very prominent (RE5).

P14 – Use audible and visible feedback to indicate functionality; add a 0.5s delay on the audio feedback and make the visual feedback very prominent

It was found during the evaluation sessions that a large number of items could be presented on the same page in the prototype application (RO1). Participants did not seem to be overwhelmed with the amount of options, some did not notice the buttons on the bottom of the screen at first but when provided with the hint (H9) were able to locate the buttons and find the one from the category they were required to select. The fact that all items were presented in the same screen certainly made it easier for children to find what they wanted than if they were distributed in two or more pages requiring navigation between different pages. It was proven to be an efficient way for preschoolers to search and browse video content (REA1). So, according to the literature, a flattened hierarchy should be used (RL25), sub-menus should be avoided (RL8), if additional pages are needed paging should replace scrolling (RL15), but ideally all items should be presented in the same screen (RL17).

P15 – Present all items in one single screen, use a flattened hierarchy with no sub-menus

The searching and browsing should be provided via pre-established categories (REA2). In order to establish these categories preschoolers' concepts of categories have to be considered (RL5), therefore it is important to involve children in this part of the design process. It was found during design sessions, for instance, that children tend to associate video content to the channel it is broadcast (RCS3), and they do not differentiate cartoons from animations (RCS4). Thus, on an EPG children's channels/brands should be included as categories and cartoons and animations should be included in the same category.

P16 – Provide searching and browsing via categories established with children during the design process

The system of classification should adopt a faceted approach (REA3) that is simpler than the hierarchical approach and provides opportunity to find items based on more than one dimension. So children may explore the interface (RO2) and refine their exploration using conjunctive Boolean searches (RO3). During evaluation sessions, participants created Boolean searches but it was not evident that they understood they were creating such conjunctive searches. Nevertheless, as discussed in Section 6.2, if children are given time and assistance they can actually accomplish Boolean searches that can be useful to the search process. Thus, interfaces for this age group should allow conjunctive Boolean searches.

P17 – Adopt a faceted approach to the system and allow conjunctive Boolean searches

The core functionality should be always visible and presented consistently following standards (RL27). The exit icon, for example, should be red (RE2). This is not a platform standard; it actually goes against the standard for iTV applications in the United Kingdom, where the red button is used to initiate the interactive services. There is a convention, however, in computer software and in general that the colour red represents exit, close, stop, and this surpassed the iTV convention. Two icons were used one to go back the other to close the application (RE1) and participants had no trouble in identifying that the buttons to go back or close were the red ones; red then may be used as an easy escape route (RL36).

P18 – Present core functionality consistently following standards, e.g. a red button/icon should be used as an escape route

The prototype interface was simple (RL32) and minimum training was needed (RL31), with some scaffold and guidance (RL9) provided through the tutorial (RE4) and hints children were able to use the system without assistance (RL30). A help section was always accessible via the same button (RL33), instructions were age-appropriate, easy to comprehend and remember (RL34), and divided into two small segments (RE3). The help button was used during evaluation sessions but it was not clear which type of assistance participants were expecting to find. The help segment was misinterpreted by some participants; it would then be useful for a character to explain the interaction process or for it to be totally different from the actual interface so that it is clear to users that the buttons displayed during the help segment are not active because they are illustrating how the interaction works, but are not part of the interface itself.

P19 – The interaction should be simple so minimum training is needed
P20 – Provide scaffold and guidance through tutorial and help section
P21 – Make the help section always accessible via the same button, containing age-appropriate instructions, divided into small segments, easy to comprehend and remember and illustrated differently from the actual interface

During the evaluation sessions there was not enough evidence that participants understood the ‘favourites’ feature of the prototype, but children talked about their favourite programmes and asked to watch videos they had already seen, this indicated a need for such feature. This feature could be useful to allow children to rate video content indicating their preferences (REA4), supporting in some way customization and personalization (RL39) of the interface, and to provide opportunity for repetition (RL38). The initial video loops did present opportunity for repetition; they were removed from the prototype following a requirement (RE7), and it does seem sensible for users to repeat only the content they choose to, not every single video they access.

P22 – Allow children to indicate their favourites
P23 – The interface should support customization and personalization
P24 – Provide opportunity for repetition

Only children's video content was included in the prototype (REA5), but there was an issue during an evaluation session in which a participant said he was too scared with the video content presented. Therefore it is important to select content specifically to the exact age group of the target audience.

P25 – All content presented should be aimed at the exact age group of the target audience

A very subtle humour was added to the interface (RL11) that apparently had no positive or negative impact on the interaction process. There were, however, funny videos available and some participants did laugh and seem to have fun while watching them. This indicates that preschool children do enjoy humour if added appropriately (RL28), it does not necessarily have to be added to the interface itself, which can sometimes interfere with the navigation, it can be added to the content instead.

P26 – Consider adding humour to the interface content

During several design and evaluation sessions it was found that children could easily relate screenshots to video content. For this reason screenshots may represent video content on interfaces that provide access to video (RCS2).

P27 – Video content may be represented by screenshots

There was an initial requirement that regional videos should be more easily accessible than imported videos (REA6). It was decided, however, not to follow such requirement but to choose for the tasks to be performed during evaluation sessions videos that were found popular with children during previously stages of the research. Some of the popular content was American but participants in Brazil and in the UK were very familiar with. The intention of the evaluation sessions was, however, to ask participants to find videos they would recognise as easily as possible and this sometimes meant favouring imported content. Nevertheless, it is certainly better for an EPG application to present first the content produced locally that will most possibly reflect the users' reality, but for this to become a principle it would have to be tested.

During the evaluation sessions it was found that some participants tried to interact with the greyed inactive buttons; it could be probably interesting to have a form of short audio feedback that would be the same for all inactive buttons to gives user the idea that the button is inactive and cannot be turned on. This, however, was not tested so it is not clear if participants would understand the audio feedback for the inactive buttons or if the sound could become annoying instead of useful. For this reason, this idea did not evolve into a principle.

Participants were allowed to control the video (RE8), pause it and play it back if they wished. The feature, however, was not recognised by any participant. This could mean participants were not familiar with the icon or did not notice it on the screen. It is not evident if users this age would like to have the option to pause a video or if they would prefer going back to the menu instead. Thus, additional testing is needed for this requirement to become a design principle.

6.4 Techniques for Design and Evaluation

Existing techniques for design and evaluation were used in the work described in this thesis such as low-tech prototyping with children and expert evaluation. Moreover, novel card sorting techniques were developed to contribute to the information architecture of the system. Most existing methods used were refined and the card sorting activities developed, these are briefly reviewed in this session and principles to assist on the use of the techniques presented. In addition, a series of guidelines for design and evaluation sessions with children were transformed into requirements for working with children (Chapter 4) and followed during the sessions conducted as part of the work described in this thesis. These requirements are supported by the findings from the sessions and transformed into principles for working with preschoolers.

The first method used involving children to contribute to the design of the prototype application was observation. It was, however, the only technique that did not require any adjustment or refinement. Children were observed in the nursery setting and during this exploratory study some initial requirements were elucidated.

The following stage of the research included the definition of the categories for the system, in which it was necessary to create card sorting techniques that could be used with preschoolers. Three techniques were developed and tested. The closed card sorting in which children were presented with a screenshot and two boxes with pre-established categories and were asked to post the screenshot in the box they found most appropriate. The open card sorting in which children were asked to put together screenshots they believed were the same type or kind. And the match-to-sample in which children were given a sticker with a screenshot and asked to paste it on the group of screenshots they thought it would be most appropriate. It was found that the closed card sorting could be very useful to determine when a category is well understood. The open card sorting needs to be further developed because it probably demands more of participants' time. The match-to-sample was found useful when combined with the closed card sorting to define and refine categories, so more than one type of card sorting activity should be conducted with preschoolers to assist structuring the information architecture of a system (RCS5). The prototype incorporated the categories that emerged from these activities and during evaluation sessions it was found that most categories were well understood by participants, facilitating the search process. Therefore, it is believed that if such techniques are used during the design process conceptual structures may be elicited from participants reflecting users' view of the content. Further work is needed to investigate the impact of the card sorting activities in the design of menu entries and/or headings for technology aimed at young children.

P28 – Use card sorting activities to contribute to the information architecture of a system, ideally more than one type of activity should be conducted with preschoolers (e.g. when taken an informant approach it is useful to combine the closed card sorting with the match-to-sample)

Another design activity conducted during the research reported in this thesis was the low-tech prototyping, based on Scaife and Rogers (1999) suggestions for low-tech prototype sessions with children, and ideas from Guha et al. (2004) to work with young children. In the sessions children were given laminated images, glue and crayons and asked to make their own version of the prototype. It was found during the process that some children talked about the interaction manipulating the images against the background. For this reason it is recommended to use more interactive materials (RLTP1) like Velcro instead of glue, so that participants can elaborate on ideas and move the elements on their prototypes while demonstrating the interaction process. It would also be good to analyse these data in detail, so it is also suggested to video record the sessions (RLTP2).

P29 – Use interactive materials for low-tech prototyping and video record the sessions if possible

Experts were asked to analyse a version of the prototype application in order to suggest improvements for its redesign. Two methods were used in this process. First experts were asked to conduct a cognitive walkthrough (Wharton et al., 1994), in which they performed a task and checked for each step how easy it would be for a new user to accomplish the task. Then they were asked to answer some questions, based on the structured expert evaluation method (Baauw et al., 2005), with their opinion about the system. It was found that the two methods complemented each other. The walkthrough was useful to provide structure for experts to explore the prototype and along with the structured expert evaluation certainly contributed to a holistic analysis of the prototype application.

P30 – Cognitive walkthroughs may be complemented by the structured expert evaluation method (SEEM) to contribute to a holistic analysis of the system by experts

There were some problems and concerns raised by experts with no clear solutions, for this reason it was decided to conduct prototype adjustments sessions with children. It was found during the study that prototype adjustments sessions, a simpler version of high tech prototyping session or agile development, can be useful to redefine icons and verify if there are any major issues during the interaction process that can be solved before the actual evaluation sessions. The high-tech prototype should be developed as flexible as needed so that adjustments can be made instantly during the evaluation session.

P31 – Prototype adjustment sessions may be an option to the high tech prototype session or agile development by providing opportunity for participants to instantly amend the prototype

During the evaluation sessions it was found very important to offer the opportunity for participants to interact with alternative input devices to accommodate children's varying abilities and fine motor skills (P2). In addition, the hints given also played an important role preventing participants to become stressed when struggling to accomplish tasks with no assistance. Thus, it is recommended when working with three and four year olds to provide some flexibility during the sessions, through alternative devices and hints, for them to feel as comfortable and confident as possible using the devices and interacting with the system.

P32 – Evaluation sessions should be flexible; offer alternative input devices for interaction and hints

A problem occurred during the evaluation sessions regarding the surface provided for participants to rest the remote control. The surface was provided based on a requirement (RL42), but it was not rigid enough. For this reason, it is recommended to provide a surface that is stable and rigid, so when rested the remote does not interfere with the navigation in any unexpected way (e.g. by striking keys on the keyboard or the touchpad). In addition, it is important to note that even rested on the surface the remote should be able to communicate with the IR receiver.

P33 – During evaluation sessions, provide a stable and rigid surface for participants to be able to rest the remote with no interference on the interaction

An adapted version of the Fun Toolkit (Read and MacFarlane, 2006) was used to gather children's opinion about the input devices they interacted with during the evaluation sessions. The materials used in the adapted version seemed to work very well with preschoolers. In the Smileyometer, instead of ticking boxes with smile faces they chose smile stickers to paste next to each device. Participants needed the words from the Again-Again table to be read to and then they could colour their answer about which device they would like to use again. The ranking system also worked well using a rubber stamp to give a star to the device they thought was the best. Children seemed to enjoy filling out the Fun Toolkit survey, the data collected, however, could not be triangulated. It was not evident if participants were not able to rate their experience with a specific device and compare the two devices, or if they wanted to balance their responses. Nevertheless, it was noticed that participants' degree of preference for each device was related to their experience with the device. Therefore the method can be used with preschool children to rate their experiences after the interaction, but the data should be gathered in different forms (Smileyometer, Again-Again Table and Ranking System), and results should be carefully calculated so that every form has similar weight, then the findings can be compared with video analysis and other indicators of the overall user experience.

P34 – An adapted version of the Fun Toolkit may be used with young children to rate their experiences after interaction, but the data should be gathered in different forms and findings compared with video analysis

A series of requirements for working with children derived through analysis of guidelines for design and evaluation sessions and were followed during the studies conducted as part of the work described in this thesis. Some of the requirements evolved into design principles following the findings from the sessions.

There was a requirement that indicated that design and evaluation sessions should be preferably conducted individually (RL4). Most sessions were conducted individually and did work well, but the low-tech prototype session was carried out in a group and presented no problems. Thus, as long as the child to adult ratio is low (RL40), so every participant can be given attention and/or assistance when needed, it is adequate to involve more than one participant in a session at a time.

P35 – Design and evaluation sessions may be conducted with a group of participants as long as the child to adult ratio is low

It was found that when the sessions were conducted in a quiet room, children were more focused on the activity, therefore it is recommended to conduct design and evaluation sessions in a quiet room (RCS1) instead of using a corner of the nursery setting.

P36 – Conduct design and evaluation sessions in a quiet room

The 30 minute sessions were found adequate for most participants, but some participants, especially the younger ones, were found to be fidgety at the end of the session. Thus, design and evaluation sessions should be brief (RL7) and the researcher should be attentive to interrupt the session in case a participant is thought to be tired or fidgety at any point.

P37 – Make design and evaluation sessions brief

Participants did seem to enjoy being complimented following the sessions (RL10) and offered a certificate and stickers to thank them for their participation. The feedback provided (RL41), by telling participants they were such great helpers, was also found important after the session and during some tasks when participants were displaying signs of lack of confidence.

P38 – Provide participants feedback and compliment them after the sessions and during the tasks if necessary

6.5 Conclusions

In this chapter findings from previous chapters were discussed and provided basis for the framework of children's interactions with iTV applications, illustrated in Figure 84. Design principles to assist on the design of technology for young children were produced. And finally, techniques for design and evaluation were reviewed. A list of the 38 principles, for input devices, iTV interfaces and for working with young children are presented in the Appendix P. The next chapter will provide a brief summary of the research findings, contributions of the thesis and ideas for future work.

Chapter 7. Conclusions and Future Work

7.1 Introduction

This chapter includes a brief summary of the thesis by first revisiting the research approach (Section 7.2) and its limitations (Section 7.3), then by summarising the major contributions (Section 7.4). In Section 7.5 some of the possibilities for future research are presented and the chapter and thesis concludes in Section 7.6.

7.2 Research Approach

The aims of the work in this thesis were to:

- Analyse children's interactions with an interactive television application and illustrate them in a framework so as to further the understanding of the way preschoolers interact with television.

- Contribute design principles for preschool interactive television.

- Refine methods and add to the knowledge of design and evaluation techniques involving young children.

The aims were met through design and evaluation sessions, and the major findings are discussed in Chapter 6.

Behind this work was the conviction that young children are able to interact with interactive television interfaces on their own using different input devices such as the remote control and mouse. The empirical work carried out supports this claim.

The research was inspired by educational design research and involved the design and evaluation of an interactive TV application in which children participated actively as informants. The approach taken worked well, but presented some limitations acknowledged in the following section.

7.3 Limitations of the Research

The research has some limitations. The framework is aimed at offering both explanation and understanding. Isolated variables were not emphasized, but studied as integral and meaningful phenomena (Akker et al., 2006). Statistics were used to calculate and present some findings, but should not be used for generalization from sample to population. As most design research, the study presented in this thesis "does not strive towards context-free generalizations" (Plomp, 2007, p. 16). Nevertheless, the conceptual framework gives other researchers the opportunity to make analytical generalizations (McKenney et al., 2006).

Design principles were developed to provide some structure and support for design of interactive television applications for young children: they were based on requirements associated with design choices and analysis of children's interactions so should be used as heuristic statements because "they provide guidance and direction, but do not give 'certainties'" (Plomp, 2007, p.22).

Refined techniques for design and evaluation sessions should also provide a basis for working with young children and may be adapted for different contexts and situations.

7.4 Contributions of the Thesis

This thesis presents the following major results:

- It identifies the main issues that influenced children's interactions with iTV, summarised in the framework in Section 6.2.1.

- It presents design principles in Section 6.3 developed through the requirements identified in Chapter 4 compared to results from the evaluation sessions reported in Chapter 5.

- Through lessons learnt, especially during the empirical work, the thesis offers refined and novel techniques to work with preschoolers in design and evaluation sessions, reviewed in Section 6.4.

7.5 Future Research

The scope of the research was limited to three and four year old users, in Brazil and the United Kingdom, testing an electronic programme guide using the remote control and mouse as input devices.

One important research agenda topic is to broaden the scope by evaluating the framework, design principles and methods with other participants in different contexts, interacting with other interfaces and applications using different input devices.

7.5.1 Participants and Contexts

The framework of children's interactions with iTV resulted from design and evaluation sessions with a small number of participants in Brazil and United Kingdom. Further research could be conducted to test whether the same factors would influence the process of interaction among a larger sample. It would also be interesting to analyse the implications of the findings to different contexts, conducting sessions in other countries and settings, such as in nurseries with different curricula and in home environments.

The design principles presented derived from requirements that emerged from design activities, and a review of existing systems and literature that included assumptions about the users. It was expected that users had characteristics and abilities according to a ‘normal’ development for their age group. Further research in this area is expected to include children with a different range of abilities/disabilities.

The research was conducted with three and four year olds: further work needs to look at the impact of the findings for a wider age group. It is possible that not only older children but adults too can benefit from technologies developed specially for preschoolers in unexpected ways (Plowman and Stephen, 2003).

7.5.2 Interfaces and Applications

Interactive television provides potential for several different types of applications including those for educational and entertainment purposes. The principles presented in this thesis do not provide a complete specification for an iTV application for young children. Nevertheless, if they are implemented it is believed that a usable application would be produced, providing a good user experience. These principles were developed based on an electronic programme guide. Most may also be useful to the design and evaluation of different applications, but further work needs to be done to confirm to what extent the findings are appropriate to other interfaces.

Techniques for working with young children in design and evaluation sessions were refined and the card sorting activities developed (described in Section 4.4) were found to be very useful in involving children in the development of the information architecture of a system in order to reflect their view of the content. Further research needs to look at using such methods to effectively organize and label functionality and content of other types of interactive systems, including websites and games.

7.5.3 Input Devices

The research was limited to traditional input devices such as the remote control, mouse, keyboard and touch pad. Further work in this area is expected to include the use of touch screens as well as tangible technology for interaction with the iTV application that may assist users with limited motor skills to interact with the television.

The fast pace of technological change reinforces the need for further studies focusing on other input devices such as motion controlled applications like Microsoft Kinect[7] and Canesta's 3DTV[8]. These technologies enable interaction via gestures, eliminating the need for a separate input device. A camera integrated with the game console (XBOX 360) or set top box identifies the user, eliminates background images and objects, and accepts commands through basic gestures.

Some findings reported in this thesis could certainly be utilized by these new technologies. The card sorting methods, for instance, could be used to create an electronic programme guide for preschoolers to interact with gestures and movements. Further studies, however, need to look at the appropriateness of the principles to such contexts. Additionally, it would be interesting to investigate if the same factors affect children's interactions when using a different or no input device.

7.6 Concluding Remarks

Due to the interdisciplinary nature of the study, findings presented in this thesis may offer different contributions and possibilities to different communities.

---

[7] *Introducing Kinect for Xbox 360* [Online]. Available: http://www.xbox.com/en-GB/kinect/ [Accessed 06/08/2010 2010].

[8] *Canesta: bringing breakthrough low-cost 3D vision to everyday devices* [Online]. Available: http://canesta.com/ [Accessed 06/08/2010 2010].

To the IDC field the main contribution was the development and refinement of techniques and principles for design and evaluation sessions involving young children. For this community it may be interesting to investigate the applicability of the card sorting method to define other types of information architecture of a system, such as headings and menu entries. It would also be worth exploring the suitability of the design principles for different interfaces, applications and input devices, aimed at other age groups and/or for users in different countries.

To the iTV community the research provides an insight into electronic programme guides for preschoolers, an overview of the main issues that impinge on children's interactions and identification of design principles for interactive television for young users. One issue to be examined would be whether a similar EPG could offer adult users a good experience. It would also be interesting to investigate other issues that could affect interaction as well as further analyse each aspect that makes up the framework, individually and in detail, with a larger and more diverse sample. Additionally, the design principles for interactive television for children could be tested in different interfaces and applications.

Other communities may find in this exploratory study a range of research questions for interdisciplinary enquiry. Developmental psychologists may be interested in the way preschoolers recognize and categorize audiovisual content. Media studies researchers may look at access to, and use and effects of interactive television. The meaning of iTV for children and the impact of the digital divide are further issues for investigation by this discipline. Moreover, the influence of the findings for different types of content, including educational and entertainment, could be explored, to see whether the principles could provide the means to enhance the potential of iTV applications to support learning and increase fun.

The research presented may be relevant to a variety of fields and disciplines, but it is particularly pertinent to the iTV community. There is very little in the literature about interactive television for children. We believe that the work described in this thesis may be a step forward in the process of understanding children's interactions with iTV applications, providing a foundation for future studies in this field.

Publications Relating to this Thesis

JOLY, A.V., PEMBERTON, L. & GRIFFITHS, R. 2009. Card Sorting Activities with Preschool Children. *In*: BLACKWELL, A.F., ed. HCI2009. Cambridge, UK. British Computer Society, pp. 204-213.

JOLY, A. V. 2009. Designing iTV Interfaces for Preschool Children. *In:* DONOSO, V., GEERTS, D., CESAR, P. & GROFF, D. D., eds. EuroiTV. Leuven, Belgium. Katholieke Universiteit Leuven, pp. 160-163.

JOLY, A.V. & SHIMABUKURU, R. 2009. Design de Interfaces para TV Digital Interativa destinadas a Crianças em Idade Pré-escolar. *In*: NUNES, P. (ed.) *Mídias Digitais & Interatividade*. João Pessoa, Brazil: EDUFPB.

JOLY, A. V. 2008. Design Methods for Child-Centred Interactive Television. *In*: LUGMAYR, A., KEMPER, S., OBRIST, M., MIRLACHIER, T. & TSCHELIGI, M., eds. Euro iTV. Salzburg, Austria. TICSP, pp. 47-48.

JOLY, A.V., PEMBERTON, L. & GRIFFITHS, R. 2008. Electronic Programme Guide Design for Preschool Children. *In*: TSCHELIGI, M., OBRIST, M. & LUGMAYR, A., eds. EuroITV. Salzburg, Austria. Springer, pp. 263-267.

JOLY, A.V. 2007. From data collection to the theoretical framework: using grounded theory to design web and iTV interfaces for pre-school children. *CARS: Chici Annual Research Symposium*. Preston, UK.

JOLY, A.V. 2007. Interactive Cross-Platform Environments for Young Children. *In*: LUGMAYR, A. & GOLEBIOWSKI, P., eds. Euro iTV. Amsterdam, The Netherlands. TICSP, pp. 299-303.

JOLY, A.V. 2007 From passive viewers to content producers: enabling pre-literate children to interact with Internet and digital TV applications on their own. *Women in Computing Research: the Hopper Colloquium*. London, UK. BCS.

JOLY, A.V. 2007. Design and Evaluation of Interactive Cross-Platform Applications for Pre-literate Children. *In:* BEKKER, T., ROBERTSON, J. & SKOV, M. B., eds. International Conference on Interaction Design and Children. Aalborg, Denmark. ACM, pp. 185-188.

Glossary

Again-Again Table		A tool, part of the Fun Toolkit, used to determine children's preferences by asking them to indicate if they would like to use the technology again.
Application		Software developed to perform a particular set of functions.
Co-operative Enquiry		Method developed by the researchers at the University of Maryland (USA), to involve children as design partners in all stages of the design process.
Electronic Programme Guide	EPG	Refers to software on the set top box or digital television that shows on-screen listing of videos available.
Fun Toolkit		Method developed by researchers at the University of Preston (UK), to measure children's technology preferences.
Interactive Television	iTV	Programmes, applications and services that the user interacts with using a device connected or directly linked to the TV screen.
Ranking System		A tool, part of the Fun Toolkit, used to rank different technology or activities.

Return Channel		Channel used to establish a two-way communication link with the service provider.
Smileyometer		A tool, part of the Fun Toolkit, used to measure children's preferences based on smiley faces.
Set top Box	STB	A receiver that processes digital signals from the source and decodes them to display on the television.
Video on Demand	VOD	Access to specific videos at the request of the user.
Zap		The prototype designed and evaluated during the studies reported in this thesis.

Appendices

Appendix A: Complete List of Requirements

Requirements from literature (RL)

RL1 – Input devices should have an appropriate size/dimension

RL2 – Explore the use of coloured buttons

RL3 – Provide flexibility

RL4 – Favour design and evaluation sessions in which activities are conducted individually

RL5 – Consider preschoolers' concept of categories

RL6 – Avoid text

RL7 – Design and evaluation sessions should be brief

RL8 – Avoid sub-menus

RL9 – Provide scaffold and guidance

RL10 – Compliment participants during design and evaluation sessions

RL11 – Add humour

RL12 – All mouse buttons should have the same functionality

RL13 – A limited number of keys should be used in the remote control

RL14 – The remote control arrow keys should be used for choosing and the OK key to confirm the choice during navigation

RL15 – Paging should replace scrolling

RL16 – Large and prominent buttons should be used for paging

RL17 – Decrease the number of categories, when possible, to fit content in only one page

RL18 – Place icons close to each other but distanced enough to compensate inaccuracy in targeting

RL19 – On-screen selectors should be prominent

RL20 – Cursor should move logically and its initial position should be close to important links

RL21 – Each icon should have at least 64 pixels diameter

RL22 – Resize cursor activation area to enhance target acquisition

RL23 – Use icons and metaphors associated with simple words

RL24 – Clickable icons should look clickable

RL25 – Use a flattened hierarchy

RL26 – Place important icons in the middle of the page

RL27 – Make core functionality always visible and present it consistently following platform standards

RL28 – Randomness and humour should be added appropriately

RL29 – The system should be flexible and efficient

RL30 – Children should be able to use the system without adult assistance

RL31 – Minimum or no training should be needed

RL32 – Make the interface simple

RL33 – Provide help and make it always accessible via the same button

RL34 – Instructions should be age-appropriate, easy to comprehend and remember

RL35 – Provide audible and visible feedback by adding animation and audio rollovers indicating functionality

RL36 – Provide an easy escape route and present its icon on screen at all times

RL37 – Promote participation

RL38 – Provide opportunity for repetition

RL39 – Support customization and personalization

RL40 – Maintain the child to adult ratio low during sessions with children

RL41 – Provide feedback for participants' actions

RL42 – Provide a surface for participants to place the remote control

Requirements from existing applications (REA)

REA1 – Allow users to search and browse video content

REA2 – Provide searching and browsing via pre-established categories

REA3 – The system of classification should adopt a faceted approach

REA4 – Allow children to rate video content indicating their preferences

REA5 – Only children's video content should be included in the application

REA6 – Regional videos should be more easily accessible than imported videos

REA7 – Video recommendations should take into account the time of the day

Requirements from observing children (RO)

RO1 – A large number of options should be provided

RO2 – Provide room for exploration and experimentation

RO3 – Allow conjunctive Boolean searches

RO4 – Enable interaction via keyboard

Requirements from card sorting activities (RCS)

RCS1 – Design and evaluation sessions should be conducted in a quiet room

RCS2 – Screenshots should represent video content on the interface

RCS3 – Children’s channels should be included as EPG categories

RCS4 – Cartoons and animations should be included in the same EPG category

RCS5 – More than one type of card sorting activity should be conducted with preschoolers to assist structuring the information architecture of a system

RCS6 – Categories should be broad

RCS7 – There should be an overlap of categories

Requirements from low-tech prototyping sessions (RLTP)

RLTP1 – During low-tech prototyping sessions interactive materials should be used

RLTP2 – Low-tech prototype sessions should be videotaped if possible

Requirements from experts’ evaluations (RE)

RE1 – Two different icons should be presented, one to go back the other to exit the application

RE2 – The exit icon should be red

RE3 – Instructions and help section should be divided into small segments

RE4 – Tutorial should be provided for inexperienced users

RE5 – Visual feedback should be prominent

RE6 –0.5 second delay should be added to audio feedback

RE7 – Video loops should be removed

RE8 – Users should be allowed to control the video

Requirements from prototype adjustment session (RPA)

RPA1 – Enable interaction via touch pad

Appendix B: Parental Information Letter

Parental Information Sheet

Your child is invited to participate in a study of how children interact with digital TV and computers. Your child's participation will help us understand what children think of interactive TV and the Web and how we can design better for them. The study will be undertaken by a University of Brighton researcher.

The study will be conducted in three stages: First the researcher will observe children in the nursery setting going about their daily routine to gain an understanding of their age group: all data at this stage will be recorded in form of written notes. The researcher will be at the One World Nursery three hours, twice a week, during two weeks. The second stage is called low-tech prototyping. A printed prototype will be shown to the children and they will be asked for responses. They may want to draw or use the laminated figures provided to make their improved version of the prototype. This will be one session and will take approximately 40 minutes. The data will also be recorded as written notes. The third and final stage of the study will be usability testing sessions. For this, a computer and a television will be set up in the nursery and your child will be asked to interact with some prototype material and give his/hers responses. This will take approximately 30 minutes and your child will only have access to children's content. During this stage your child will be video recorded so the data can be analysed in detail by the researcher. The videotapes will be only shown to members of the research team (Ana Joly, Lyn Pemberton and Richard Griffiths). The only exception to this is that if you give your permission, we may use portions of the videos in academic conferences or publications. They will be kept safe during the research and destroyed following the analysis.

The participation in this study is completely voluntary and you or your child may withdraw the consent to participate at any time during the process. Each stage of the research will be briefly explained to your child who may decide if he/she wants to participate and will also be free to withdrawal, stop prototyping or stop testing whenever he/she wants. You may also ask for the researcher or nursery staff to remove your child from the study at any time.

If there is any concern regarding the conduct of this study or if you would like to have access to the results of the research, you may contact Ana tel. 07716177112, Lyn Pemberton tel. 01273 642476 or Richard Griffiths tel. 01273 642477.

Please note:

This is a test of our interactive material – not of your child!

Your child's personal details will not be used for any purpose other than this study. No details relating to your child will be passed to any third party, and the published results of this study will have all data anonymised wherever possible.

We will take care to ensure your child's safety whilst participating in the study.

Appendix C: Consent Form

Consent Form

1) I accept and agree for myself and on behalf of my child to take part in this study of Interactive Cross-platform Environments for Pre-school Children.

2) The purpose of the study was explained to my satisfaction. I understand the principles and procedures fully.

3) I am aware that my child will be required to interact with the television and computer and will be asked to give his/her opinion about hi-tech and low-tech prototypes presented.

4) I understand that my child is free to withdraw from the investigation at any time without giving a reason.

5) I have read the description of the study and agree that my child may participate on the terms set out above.

I consent to the use of sections of the video tapes in scientific conference presentations and journal articles ☐

Child's Name: ____

Parent or Guardian's Name: __________________________________

Signature: ____

Date: ____

Many thanks for your help and cooperation, Ana Vitoria Joly Hulshof.

Appendix D: Parental Information Letter (Card Sorting)

Parental Information Sheet

Your child is invited to participate in a study of how children interact with information retrieval systems on digital television and computers. Your child's participation will help us understand what children think of interactive TV and the Web and how we can design better for them. The study will be undertaken by a University of Brighton researcher, Ana Joly.

To identify how children categorize audiovisual content we have developed a card sorting activity in which children classify television programmes, movies and music videos. The results from this card sorting session will define categories to be implemented in a prototype information retrieval system under development. Each child will carry out a card sort task individually in which they will be presented with a picture of a piece of children's audiovisual content and will be asked to put it in the box they find most appropriate. This activity would take approximately seven minutes and will be conducted in a corner of the Nursery setting or a separate room under the supervision of a member of Nursery staff. In our preliminary trials children have found this a fun activity.

All data will be recorded in the form of written notes, children will NOT be photographed or video recorded.

The participation in this study is completely voluntary. The research will be briefly explained to your child who may decide if he/she wants to participate and will be free to withdraw whenever he/she wants.

If there is any concern regarding the conduct of this study or if you would like to have access to the results of the research, you may contact Ana tel. 07716177112, Lyn Pemberton tel. 01273 642476 or Richard Griffiths tel. 01273 642477.

Please note:

This is a test of our interactive material – not of your child!

Your child's personal details will not be used for any purpose other than this study. No details relating to your child will be passed to any third party, and the published results of this study will have all data anonymised.

We will take care to ensure your child's safety whilst participating in the study.

Appendix E: Consent Form (Card Sorting)

Consent Form

1. I accept and agree for myself and on behalf of my child to take part in this study of Audiovisual Information Retrieval System for Young Children.
2. The purpose of the study was explained to my satisfaction. I understand the principles and procedures fully.
3. I am aware that my child will be required to perform a card sorting activity.
4. I understand that my child is free to withdraw from the investigation at any time without giving a reason.
5. I have read the description of the study and agree that my child may participate on the terms set out above.

Child's Name: ____

Parent or Guardian's Name: ________________________________

Signature: ____

Date: ____

Many thanks for your help and cooperation, Ana Vitoria Joly Hulshof.

Appendix F: Triads of Screenshots and Categories (Closed Card Sorting)

Triads of screenshots and the two categories presented for the closed card sorting.

Screenshots	Expected Category	Unexpected Category
Ratatouille / Cars	Movies	Music
Bob the Builder / Dora the Explorer	Cartoons	Fairy Tales
Fraggle Rock / Tweenies	Music	Superheroes
Balamory / Tweenies (both make and do scenes)	Make and Do	Animals and Nature
Wild Show / Saving Species	Animals and Nature	Cartoons
Cinderella / Barbie Mariposa And Her Butterfly Fairy Friends	Fairy Tales	TV Shows
Ben 10 / Kimpossible	Superheroes	Around the World
Serious Amazon / George of the Jungle	Around the World	Cartoons
Teletubbies / Pocoyo	TV Shows	Superheroes
Balamory / Pingu	CBeebies (channel)	CITV (channel)
Goofy / Club House Disney	Disney (channel)	CBBC (channel)

Appendix G: Sets of Screenshots handed to Participants (Open Card Sorting)

Condition 1

1. Wild Show, Save the Species, Barbie Mariposa And Her Butterfly Fairy Friends, Cinderella
2. Balamory, Balamory (make and do scene), Tweenies (music scene), Fraggle Rock
3. Goofy, Club House Disney, Pingu, Pocoyó
4. Ben 10, Dora the Explorer, Kimpossible, Powerpuff Girls
5. Cars, Bob the Builder, Ratatouille, Teletubbies

Condition 2

1. Wild Show, Save the Species, Cinderella, Balamory, Balamory (make and do scene)
2. Barbie Mariposa And Her Butterfly Fairy Friends, Tweenies (music scene), Fraggle Rock, Goofy, Club House Disney
3. Pocoyó, Pingu, Ben 10, Dora the Explorer, Powerpuff Girls
4. Kimpossible, Cars, Bob the Builder, Ratatouille, Teletubbies

Appendix H: Screenshots and Categories Presented (Match-to-Sample)

Condition 1

Screenshots	Best Exemplar's Category	Also Belongs to the Categories	Not a Member of the Categories
Bob the Builder	Cartoons and Animations	-	TV Shows, Movies, Superheroes, Make and Do, Fairy Tales
Ratatouille	Movies	Cartoons and Animations	Make and Do, Music, Superheroes, Fairy Tales
Ben 10	Superheroes	Cartoons and Animations	Make and Do, Movies, Fairy Tales, Music

Condition 2

Screenshots	Best Exemplar's Category	Also Belongs to the Categories	Not a Member of the Categories
Teletubbies	TV Shows	-	Cartoons and Animations, Fairy Tales, Make and Do, Movies, Superheroes
Balamory (make and do scene)	Make and Do	TV Shows	Superheroes, Movies, Fairy Tales, Cartoons and Animations
Cinderella	Fairy Tales	Cartoons and Animations, Movies	Superhero, Make and Do, Music

Condition 3

Screenshots	**Best Exemplar's Category**	**Also Belongs to the Categories**	**Not a Member of the Categories**
Tweenies (music scene)	Music	-	Make and Do, Movies, Cartoons and Animations, Superheroes, Fairy Tales
Balamory	TV Shows	-	Cartoons and Animations, Fairy Tales, Superheroes, Movies, Make and Do
Goofy	Cartoons and Animations	-	Make and Do, Music, Movies, Superheroes, Fairy Tales

Appendix I: Expert Evaluation Form

University of Brighton

School of Computing Mathematical and Information Sciences

Prepared by Ana Vitoria Joly

Research Supervisors: Dr Lyn Pemberton and Richard Griffiths

June, 2008.

Dear Expert,

Purple is an Electronic Programme Guide made for Interactive TV and Web to assist young children (ages 3 and 4 years old) to find audiovisual content using the remote control and the mouse respectively.

We would love to hear your opinion to improve this system so it can be shown to children, in the following stage of the project, with a reduced number of navigational problems and we can then focus our attention in the analyses of their interactions which is our main aim.

Please note that the prototype is not complete there is only one video available.

We appreciate if you could follow the cognitive walkthrough below and then complete the worksheet and the expert evaluation. This should take approximately 15 minutes. Please feel free to give us more feedback as well, of any kind.

The data collected in this form might be used in publications but it will be always anonymized. We will provide you a digital copy of any publication and the final thesis.

Thank you very much for your help.

Your Details:

Company:

Contact Information:

Field of Expertise:

Children☐ Technology for Children ☐ Human-Computer Interaction☐

Digital Television☐ Other: ___________________

Years of Experience in the Field:

Less than1 ☐ 1 to 5 ☐ 5 to 10 ☐

10 to 15 ☐ More than 15 ☐

Set up Information:

Please access the prototype from the webpage: http://www.digiduo.com/purple

You can download the prototype or access it on-line. We strongly recommend to download and save the 'purple.zip' file on your computer, extract all files and then open the 'purple.exe' file. You will find detailed instruction on the page above.

If you choose to evaluate the prototype on-line you will need Flash Player. In case you don't have it installed in your computer please download it from the link below:

http://www.adobe.com/shockwave/download/flash/trigger/en/3/index.html

Once you have opened the 'purple.exe' file please follow the walkthrough on the next page and complete the walkthrough worksheet. After that please fill in the Structured Expert Evaluation with your opinion about the system.

Please note that audio may assist the navigation.

Walkthrough:

Task:	Find “Bob the Builder” music video.	
Step	**Description:**	**Goal:**
1	Click on the “Cartoons” button.	Refine Search
2	Click on the “Cbeebies” button.	Refine Search
3	Click on the “Music” button.	Refine Search
4	Click on the “Bob the Builder” thumbnail. on the upper screen.	Select Music Video
5	Watch a segment of the episode.	Watch the music video
6	Click on the “Exit”. The green button on the top of the screen.	Return to the Main Menu
7	End of the Task	

Please note that preschool children will be able to perform this same task using the mouse, or the remote control. With the remote control the arrow keys would be used to choose a button, the SELECT button would be used as the click (steps 1-4) and the GREEN button would be used to exit (step6).

Cognitive Walkthrough Worksheet:

For each step on the table above please check if:

A. Will the users be trying to produce whatever effect the action has?

B. Will users be able to notice that the correct action is available?

C. Once users find the correct action at the interface, will they know that it is the right one for the effect they are trying to produce?

D. After the action is taken, will users understand the feedback they get?

Please answer for each step **yes** or **no** from questions A to D and insert your comments and suggestions on the following columns.

Step	A	B	C	D	Comments/ Issues	Suggestions
1						
2						
3						
4						
5						
6						

Please fill in the Structured Expert Evaluation (SEEM) with your opinion about the overall system. (Based on Baauw, Bekker and Barendregt, 2005)

1) Do children understand the system?

2) Do children know what to do in order to interact with the system?

3) Are children able to perform physical actions easily?

4) Can children perceive the feedback? This includes feedback (if any) from both correct and wrong actions, and whether children can click to stop the feedback.

5) Do children understand the feedback? This holds for both visual and auditory feedback from correct and wrong actions.

6) Will children like the search process?

7) Is the level of difficulty okay for young children (3 years old)?

8) Are the navigation possibilities and the exits clear?

9) Are the icons in the interface meaningful? Are there any items that will cause problems?

10) Is the interface tuned to children's universe?

11) Is the help session available easy to access?

12) Is the help information appropriate and comprehensible for young children?

Additional comments:

Please save and send me this evaluation.

Thank you very much, Ana – A.V.B.Joly@brighton.ac.uk

Appendix J: Parental Information Letter (Evaluation Sessions)

Parental Information Sheet

Your child is invited to participate in the study: Interface Design for Preschool Interactive Television. Your child's participation will help us understand what children think of interactive TV and the Web and how we can design better for them. The study will be undertaken by a University of Brighton researcher, Ana Joly.

In order to analyse preschoolers' interactions we developed a prototype of an information retrieval system for children's videos in which they may interact using the remote control or mouse. This involves each child individually carrying out video searches in which they will be presented with a short tutorial and will be asked to find some videos using the prototype. Segments of videos, that last about two minutes, will be displayed including cartoons, music and films. All content is aimed for children audience and appropriate for preschoolers. At the end of the activity children will be asked to evaluate the prototype using stickers, crayons and a stamp. This activity would take no more than thirty minutes with each child and will be conducted in the Nursery. In our preliminary trials children have found this a fun activity.

Your child will be video recorded so the data can be analysed in detail by the researcher. The videotapes will be only shown to members of the research team (Ana Joly, Lyn Pemberton and Richard Griffiths). The only exception to this is that if you give your permission, we may use portions of the videos in academic conferences or publications. They will be kept safe during the research and destroyed following the analysis.

The participation in this study is completely voluntary. The research will be briefly explained to your child who may decide if he/she wants to participate and will be free to withdraw whenever he/she wants.

If there is any concern regarding the conduct of this study or if you would like to have access to the general results of the research, you may contact Ana tel. 07716177112, Lyn Pemberton tel. 01273 642476 or Richard Griffiths tel. 01273 642477.

Please note:

This is a test of our interactive material – not of your child!

Your child's personal details will not be used for any purpose other than this study. No details relating to your child will be passed to any third party, and the published results of this study will have all data anonymised wherever possible.

We will take care to ensure your child's safety whilst participating in the study.

Appendix K: Consent Form (Evaluation Sessions)

Consent Form

6. I accept and agree for myself and on behalf of my child to take part in this study of Interface Design for Preschool Interactive Television.
7. The purpose of the study was explained to my satisfaction. I understand the principles and procedures fully.
8. I am aware that my child will be required to interact with the prototype presented using the remote control or mouse and will be asked to give his/hers opinion about it.
9. I understand that my child is free to withdraw from the investigation at any time without giving a reason.
10. I have read the description of the study and agree that my child may participate on the terms set out above.

I consent the use of sections of the video tapes in scientific conference presentations and journal articles ☐

Child's Name: ____

Child's Age: _____ Years _____ Months

Parent or Guardian's Name: ______________________________________

Signature: ____

Date: ____

Appendix L: Parents' Questionnaire (Evaluation Sessions)

Brief Questionnaire

Dear Parent/Guardian,

We would also appreciate if you could complete the following questions and return it with the consent form.

1. How often does your child watch television? *Please only cross ONE box*

A. Never ☐
B. Several times per day☐
C. About once a day ☐
D. A couple of times a week ☐
E. About once a week☐
F. About a couple of times a month ☐
G. About once a month ☐
H. Less Often☐
I. Don't Know ☐

2. Does your child use the remote control when watching television?

A. Yes ☐
B. No ☐
C. Only with supervision ☐

3. How often does your child use the computer from any location (e.g. nursery, home)? *Please only cross ONE box.*

A. Never☐
B. Several times per day☐
C. About once a day ☐
D. A couple of times a week ☐
E. About once a week☐
F. About a couple of times a month ☐
G. About once a month ☐
H. Less Often☐
I. Don't Know ☐

4. Does your child use the mouse when using the computer?

A. Yes ☐
B. No ☐
C. Only with supervision ☐

Many thanks for your help and cooperation, Ana Joly.

Appendix M: Adapted Fun Toolkit

Choose one sticker for each activity.

Awful

Not very good

Good

Really good

Brilliant

Would you like to do it again?

Please colour your answer.

	
Yes	Yes
Maybe	Maybe
No	No

Which one is the best?

Please stamp your answer.

Appendix N: Conditions during Testing Sessions

Condition1

Mouse first, remote second	
Order of Tasks Brazil: 1) Cocoricó 2) Cinderela 3) TV Xuxa 4) Dora a Aventureira 5) Ben 10 6) Ratatouille	Order of Tasks UK: 1) Pingu 2) Cinderella 3) Tweenies 4) Dora the Explorer 5) Ben 10 6) Ratatouille
Evaluation form 1	

Condition2

Remote first, mouse second	
Order of Tasks Brazil: 1) Cocoricó 2) Cinderela 3) TV Xuxa 4) Dora a Aventureira 5) Ben 10 6) Ratatouille	Order of Tasks UK: 1) Pingu 2) Cinderella 3) Tweenies 4) Dora the Explorer 5) Ben 10 6) Ratatouille
Evaluation form 2	

Condition3

Mouse first, remote second	
Order of Tasks Brazil: 1) Dora a Aventureira 2) Ben 10 3) Ratatouille 4) Cocoricó 5) Cinderela 6) TV Xuxa	Order of Tasks UK: 1) Dora the Explorer 2) Ben 10 3) Ratatouille 4) Pingu 5) Cinderella 6) Tweenies
Evaluation form 1	

Condition4

Remote first, mouse second	
Order of Tasks Brazil: 7) Dora a Aventureira 8) Ben 10 9) Ratatouille 10) Cocoricó 11) Cinderela 12) TV Xuxa	Order of Tasks UK: 7) Dora the Explorer 8) Ben 10 9) Ratatouille 10) Pingu 11) Cinderella 12) Tweenies
Evaluation form 2	

Appendix O: Testing Session Script

Hi, I'm Ana and I've been working on something to help children find programmes on the TV and on the computer, but, I don't know many children's programmes so I need your help to improve it, could you help me please? Thanks. If you want to stop helping me at any time, please say and we'll stop, no problem, all right? Good.

Do you watch TV? How do you choose what to watch? How about the computer, do you use it? What do you use it for? Great.

Should we get started then? I'll show you how it works and then ask you to find some programmes, OK? Good. Because we are still working on it there are just tiny bits of the programmes to watch, we don't have the entire programmes yet, I'm sorry. (open the application and show tutorial)

Do you know programme 1? Could you please find it? Well done!

How about programme 2, do you know it? Do you think you can find it? Well done!

And programme 3? Do you know it? How do you think we can find it? Well done!

You did such a great job! May I show you another way to find programmes now? This is similar to the other one but a bit different, I'll show you how it works OK? (change to remote or mouse- show tutorial)

Do you know programme 1? Could you please find it? Well done! You don't have to watch it if you don't want. You can go back to choose another programme.

How about programme 2, do you know it? Do you think you can find it? Well done!

And programme 3? Do you know it? How do you think we can find it? Well done!

Do you think you could pause the video if you wanted?

Great! Now you've finished. Would you like to choose one video of your choice? You can use the remote or the mouse, what would you like to use?

Oh thank you very much. Just one more thing. (show evaluation forms)

We have several stickers here, brilliant, really good, good, not very good and awful. Could you please choose a sticker for this one and one for this one?

Would you like to use this one or this one again? You can colour YES, if you want to use again, MAYBE if you are not sure and NO if you don't want to use again.

The last one. Between these two please stamp the one you think was the best one?

Thank you very much, you were really helpful. Here is a certificate and you can choose a sticker. Thank you!

Appendix P: List of Design Principles

Design Principles for Input Devices

P1 – Use a one-button mouse as input device

P2 – Enable interaction via alternative input devices

P3 – Use a limited number of keys for interaction with the remote control

P4 – Explore the use of coloured buttons

P5 – Illustrate all buttons used for interaction on the screen

P6 – Enable one button press interaction

P7 – The remote control should be small enough for children to hold comfortably

Design Principles for iTV Interfaces

P8 – The system should be flexible and efficient; provide alternative ways of interaction to cater to the entire range of target users

P9 – Interfaces should be inspired by things children are familiar with; make clickable items look clickable

P10 – Use metaphors associated with labels to represent icons

P11 – Use audio feedback instead of text

P12 – Make icons with at least 64 pixels diameter, place them relatively close to each other and make the cursor activation area as big as possible without causing confusion

P13 – Place important icons in the middle of the page and make the initial position of the cursor close to them

P14 – Use audible and visible feedback to indicate functionality; add a 0.5s delay on the audio feedback and make the visual feedback very prominent

P15 – Present all items in one single screen, use a flattened hierarchy with no sub-menus

P16 – Provide searching and browsing via categories established with children during the design process

P17 – Adopt a faceted approach to the system and allow conjunctive Boolean searches

P18 – Present core functionality consistently following standards, e.g. a red button/icon should be used as an escape route

P19 – The interaction should be simple so minimum training is needed

P20 – Provide scaffold and guidance through tutorial and help section

P21 – Make the help section always accessible via the same button, containing age-appropriate instructions, divided into small segments, easy to comprehend and remember and illustrated differently from the actual interface

P22 – Allow children to indicate their favourites

P23 – The interface should support customization and personalization

P24 – Provide opportunity for repetition

P25 – All content presented should be aimed at the exact age group of the target audience

P26 – Consider adding humour to the interface content

P27 – Video content may be represented by screenshots

Techniques for Design and Evaluation

P28 – Use card sorting activities to contribute to the information architecture of a system, ideally more than one type of activity should be conducted with preschoolers (e.g. when taken an informant approach it is useful to combine the closed card sorting with the match-to-sample)

P29 – Use interactive materials for low-tech prototyping and video record the sessions if possible

P30 – Cognitive walkthroughs may be complemented by the structured expert evaluation method (SEEM) to contribute to a holistic analysis of the system by experts

P31 – Prototype adjustment sessions may be an option to the high tech prototype session or agile development by providing opportunity for participants to instantly amend the prototype

P32 – Evaluation sessions should be flexible; offer alternative input devices for interaction and hints

P33 – During evaluation sessions, provide a stable and rigid surface for participants to be able to rest the remote with no interference on the interaction

P34 – An adapted version of the Fun Toolkit may be used with young children to rate their experiences after interaction, but the data should be gathered in different forms and findings compared with video analysis

P35 – Design and evaluation sessions may be conducted with a group of participants as long as the child to adult ratio is low

P36 – Conduct design and evaluation sessions in a quiet room

P37 – Make design and evaluation sessions brief

P38 – Provide participants feedback and compliment them after the sessions and during the tasks if necessary

References

AHONEN, A., TURKKI, L., SAARIJÄRVI, M., LAHTI, M. & VIRTANEN, T. 2008. Guidelines for Designing Easy-to-Use Interactive Television Services: Experiences from the ArviD Programme. *In:* LEKAKOS, G., CHRORIANOPOULOS, K. & DOUKIDIS, G. (eds.) *Interactive Digital Television: Technologies and Applications.* London: IGI Publishing.

AKKER, J. V. D. 1999. Principles and methods of development research. *In:* AKKER, J. V. D., BRANCH, R. M., GUSTAFSON, K. & NIEVEEN, N. (eds.) *Design approaches and tools in education and training.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

AKKER, J. V. D., GRAVEMEIJER, K. & MCKENNEY, S. 2006. Introducing Educational Design Research *In:* AKKER, J. V. D., GRAVEMEIJER, K., MCKENNEY, S. & NIEVEEN, N. (eds.) *Educational Design Research.* Oxon Routledge.

ALEXANDER, A. 2004. Children and Television. *In:* NEWCOMB, H. (ed.) *Encyclopedia of Television.* Chicago, USA: Museum of Broadcast Communications.

ANTLE, A. N. 2006. Child-personas: fact or fiction? *In:* Proceedings of the 6th ACM conference on Designing Interactive systems,University Park, PA, USA. ACM, pp. 22-30.

AROYO, L., BELLEKENS, P., BJORKMAN, M., HOUBEN, G.-J., AKKERMANS, P. & KAPTEIN, A. 2007. Sensee Framework for Personalized Access to TV Content. *In:* CESAR, P., CHRORIANOPOULOS, K. & JENSEN, J. F., eds. Euro iTV,Amsterdam, The Netherlands. Springer, pp. 156-165.

BAAUW, E., BEKKER, M. M. & BARENDREGT, W. 2005. A Structured Expert Evaluation Method for the Evaluation of Children's Computer Games. *Human-Computer Interaction - INTERACT 2005.* Berlin: Springer

BAKHSHI, O. 2007. The Implications of Convergence on EPG Design: Enabling the Mediation of Content through Converged Technologies. *In:* LUGMAYR, A. & GOLEBIOWSKI, P., eds. Euro iTV,Amsterdam, The Netherlands TICSP, pp. 9-16.

BARKHUUS, L. 2009. Television on the internet: new practices, new viewers. *In:* Proceedings of the 27th international conference extended abstracts on Human factors in computing systems,Boston, MA, USA. ACM, pp. 2479-2488.

BAYDAR, N., KAĞITÇIBAŞI, Ç., KÜNTAY, A. C. & GÖKŞEN, F. 2008. Effects of an educational television program on preschoolers: Variability in benefits. *Journal of Applied Developmental Psychology,* 29**,** pp. 349-360.

BEKKER, M. M., BAAUW, E. & BARENDREGT, W. 2008. A comparison of two analytical evaluation methods for educational computer games for young children. *Cognition, Technology & Work,* 10**,** pp. 129-140.

BRIGHT, B. 2009. *Curriculum* [Online]. London. Available: http://www.barnabybright.ik.org/p_Curriculum.ikml [Accessed 17th August 2009].

BROUGHTON, V. 2006. The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings: New Information Perspectives,* 58**,** pp. 49-72.

BRUCKMAN, A. & BANDLOW, A. 2002. HCI for Kids. *In:* JACKO, J. & SEARS, A. (eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications.* NJ: Lawrence Erlbaum and Associates.

BRYANT, A. J., AKERMAN, A. & DRELL, J. 2008. Wee Wii: Preschoolers and motion-based game play. *58th Annual Convention of the International Communication Association.* Montreal, Quebec, Canada.

BUCKINGHAM, D. (ed.) 2002. *Small Screens: Television for Children,* London: MPG Books.

BUCKINGHAM, D., DAVIES, H., JONES, K. & KELLEY, P. 1999. *Children's Television in Britain,* London, British Film Institute.

BURNS, R. W. 1998. *Television: An International History of the Formative Years,* London, The Institution of Engineering and Technology.

BURTON, E. & GOULD, M. 2005. Moovl. *In:* ACM SIGGRAPH 2005 Educators program,Los Angeles, California. ACM Press, pp. 26.

CAMPOS, J. C. & MANO, A. 2006. Cognitive walkthroughs in the evaluation of user interfaces for children. *In:* Conferência Nacional em Interacção pessoa-máquina Braga, Portugal.

CAPARELLI, S. 1982. *Televisão, programas infantis e a criança,* Porto Alegre, Mercado Aberto.

CAREY, J. 1996. Winky Dink to Stargazer: Five Decades of Interactive Television. *Conference of Interactive Television.* Edinburgh.

CARMICHAEL, A., PETRIE, H., HAMILTON, F. & FREEMAN, J. 2003. The Vista Project: Broadening Access To Digital TV Electronic Programme Guides. *PsychNology Journal,* 1**,** pp. 229-241.

CARNEIRO, V. L. Q. 1999. *Castelo Rá-Tim-Bum: o educativo como entretenimento,* São Paulo, Anna Blume.

CHIASSON, S. & GUTWIN, C. 2005. Design Principles for Children's Software. Computer Science Department, University of Saskatchewan.

CHORIANOPOULOS, K. & LEKAKOS, G. 2007. Learn and play with interactive TV. *Computers in Entertainment,* 5**,** pp. 1-9.

CHRISTAKIS, D. A., ZIMMERMAN, F. J., DIGIUSEPPE, D. L. & MCCARTY, C. A. 2004. Early Television Exposure and Subsequent Attentional Problems in Children *Pediatrics,* 113**,** pp. 708-713.

CLOSE, R. 2004. Television and language development in the early years: a review of the literature. London: National Literacy Trust.

COPER, A. 1999. *The Inmates are Running the Asylum,* Indiana, Indianapolis, SAMS.

DALY-JONES, O. 2003. Case Study: The usability of electronic programme guides. *In:* GAWLINSKI, M. (ed.) *Interactive Television Production.* Oxford: Focal Press.

DAVIES, M. M. 1989. *Television is Good for Your Kids,* London, Hilary Shipman.

DAVIES, M. M. 2001. *'Dear BBC': Children, Television Storytelling and the Public Sphere,* Cambridge, Cambridge University Press.

DAVIES, M. M. & THORNHAM, H. 2007. Academic Literature Review: The future of children's television programming. Ofcom.

DEAK, G. O., PICK, A. D. & RAY, S. D. 2002. Matching and Naming Objects by Shape or Function: Age and Context Effects in Preschool Children. *Developmental Psychology,* 38**,** pp. 503-518.

DIAPER, D. 2004. Understanding Task Analysis for Human-Computer Interaction. *In:* DIAPER, D. & STANTON, N. (eds.) *The Handbook of Task Analysis for Human-Computer Interaction.* London: Lawrence Erlbaum Associates

DIGITALUK & OFCOM 2008. Switchover Progress Report Q4 2007.

DONKER, A. & REITSMA, P. 2006. Aiming and clicking in young children's use of the computer mouse. *Computer in Human Behaviour,* 23**,** pp. 2863-2874.

DRUIN, A. 2002. The Role of Children in the Design of New Technology *Behaviour and Information Technology,* 21**,** pp. 1-25.

DRUIN, A., BEDERSON, B. B., BOLTMAN, A., MIURA, A., KNOTTS-CALLAHAM, D. & PLAT, M. 1999. Children as our Technology Design Partners. *In:* DRUIN, A. (ed.) *The Design of Children's Technology.* San Francisco: Morgan Kaufmann Publishers.

DRUIN, A., BEDERSON, B. B., HOURCADE, J. P., SHERMAN, L., REVELLE, G., PLATNER, M. & WENG, S. 2001. Designing a digital library for young children. *In:* Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries,Roanoke, Virginia, United States. ACM Press, pp. 398-405.

DTV. 2010. *Cidades onde a TV digital está no ar* [Online]. São Paulo: Site Oficial da TV Digital Brasileira. Available: http://www.dtv.org.br/materias.asp?menuid=3&id=11 [Accessed 25th February 2010].

DUFFY, B. 1998. *Supporting Creativity and Imagination in the Early Years,* Buckingham, Open University Press.

DUNN, G. 1977. *The Box in the Corner: Television and the Under-fives,* London, The Macmillan Press.

ERON, L. D., HUESMANN, L. R., MONROE M. LEFKOWITZ & WALDER, L. O. 1972. Does television violence cause aggression? *American Psychologist,* 27**,** pp. 253-263.

ERONEN, L. & VUORIMAA, P. 2000. User interfaces for digital television: a navigator case study. *In:* Proceedings of the working conference on Advanced visual interfaces,Palermo, Italy. ACM, pp. 276-279.

FISCH, S. M. & TRUGLIO, R. T. 2001. *" G" is for" growing": Thirty Years of Research on Children and Sesame Street,* Mahwah, New Jersey, Lawrence Erlbaum.

FOGG, B. J., CYPHER, A., DRUIN, A., FRIEDMAN, B. & STROMMEN, E. 1999. Is ActiMates Barney ethical?: the potential good, bad, and ugly of interactive plush toys. *In:* CHI '99 extended abstracts on Human factors in computing systems,Pittsburgh, Pennsylvania. ACM, pp. 91-92.

GAWLINSKI, M. 2003. *Interactive Television Production,* Oxford, UK, Focal Press.

GENTZKOW, M. & SHAPIRO, J. M. 2008. Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study. *The Quarterly Journal of Economics,* 123**,** pp. 279-323.

GIBSON, L., NEWALL, F. & GREGOR, P. 2003. Developing a web authoring tool that promotes accessibility in children's designs. *In:* Proceeding of the 2003 conference on Interaction design and children,Preston, England. ACM Press, pp. 23-30.

GILUTZ, S. & NIELSEN, J. 2002. Usability of Websites for Children: 70 Design Guidelines. Fremont, CA, USA: Nielsen Norman Group.

GOSWAMI, U. 1998. *Cognition in Children,* Hove, Psychology Press Ltd.

GROSSMAN, T. & BALAKRISHNAN, R. 2005. The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. *In:* Proceedings of the SIGCHI conference on Human factors in computing systems,Portland, Oregon, USA. ACM, pp. 281-290.

GUHA, M. L., DRUIN, A., CHIPMAN, G., FAILS, J. A., SIMMS, S. & FARBER, A. 2004. Mixing ideas: a new technique for working with young children as design partners. *In:* Proceeding of the 2004 conference on Interaction design and children: building a community,Maryland. ACM Press, pp. 35-42.

GUNTER, B. & FURNHAM, A. 1998. *Children as Consumers: A psychological analysis of the young people's market,* London, Routledge.

HANNA, L., RISDEN, K. & ALEXANDER, K. 1997. Guidelines for usability testing with children. *Interactions,* 4**,** pp. 9-14.

HANNA, L., RISDEN, K., CZERWINSKI, M. & ALEXANDER, K. J. 1999. The Role of Usability Research in Designing Children's Computer Products. *In:* DRUIN, A. (ed.) *The Design of Children's Technology.* San Francisco: Morgan Kaufmann Publishers.

HARTSON, H. R., ANDRE, T. S. & WILLIGES, R. C. 2001. Criteria For Evaluating Usability Evaluation Methods. *International Journal of Human-Computer Interaction,* 13**,** pp. 373-410.

HAYES, D. 2008. *Anytime Playdate: Inside Preschool Entertainement Boom, How Television Became my Baby's Best Friend,* New York, Free Press.

HERIGSTAD, D. & WICHANSKY, A. 1998. Designing user interfaces for television. *In:* CHI 98 conference summary on Human factors in computing systems,Los Angeles, California, United States. ACM Press, pp.165-166.

HOME, A. 1993. *Into the Box of Delights A History of Children's Television,* London, BBC Books.

HOURCADE, J. P. 2008. *Interaction Design and Children,* Delft, The Netherlands, Now Publishers Inc.

HOURCADE, J. P., BEDERSON, B. B., DRUIN, A. & GUIMBRETIERE, F. 2004. Differences in pointing task performance between preschool children and adults using mice. *ACM Transactions on Computer-Human Interaction,* 11**,** pp. 357-386.

HOURCADE, J. P., CROWTHER, M. & HUNT, L. 2007. Does mouse size affect study and evaluation results?: a study comparing preschool children's performance with small and regular-sized mice. *In:* Proceedings of the 6th international conference on Interaction design and children,Aalborg, Denmark. ACM, pp. 109-116.

HSIEH, F. Y. 1989. Sample Size Tables for Logistic Regression. *Statistics in Medicine,* 8**,** pp. 795-802.

HSU, S. H., WEN, M.-H., LIN, H.-C., LEE, C.-C. & LEE, C.-H. 2007. AIMED- A Personalized TV Recommendation System. *In:* CESAR, P., CHRORIANOPOULOS, K. & JENSEN, J. F., eds. Euro iTV,Amsterdam, The Netherlands. Springer, pp. 166-174.

HUDSON, W. 2005. Playing your cards right: getting the most from card sorting for navigation design. *Interactions,* 12**,** pp. 56-58.

HUNTER, E. J. 2002. *Classification Made Simple,* Aldershot, Ashgate Publishing Company.

HUTCHINSON, H. 2005. *Children’s interface design for searching and browsing.* Doctor of Philosophy, University of Maryland.

HUTCHINSON, H. B., BEDERSON, B. B. & DRUIN, A. 2006. The evolution of the international children's digital library searching and browsing interface. *In:* Proceeding of the 2006 conference on Interaction design and children,Tampere, Finland. ACM Press, pp. 105-112.

HYND, A. 2006. *Evaluating four and five-year old children's responses to interactive television programs.* Doctor of Philosophy, Murdoch University.

IATRINO, A. & MODEO, S. 2007. EPG-Board a Social Application for the OmegaBox Media Center. *In:* CESAR, P., CHRORIANOPOULOS, K. & JENSEN, J. F., eds. Euro iTV,Amsterdam, The Netherlands. Springer, pp. 31-36.

IBGE 2009. Acesso à Internet e posse de telefone móvel celular para uso pessoal 2008. Instituto Brasileiro de Geografia Estatística.

ISO-13407 1999. Human centred design processes for interactive systems.

JENSEN, J. F. 2008. Interactive Television - A Brief Media History. *In:* TSCHELIGI, M., OBRIST, M. & LUGMAYR, A., eds. Euro iTV,Salzburg, Austria. Springer, pp. 1-10.

JONES, T. 1991. An Empirical-Study of Childrens Use of Computer Pointing Devices. *Journal of Educational Computing Research,* 7**,** pp. 61-76.

KAHN, P. H., FRIEDMAN, B., PÉREZ-GRANADOS, D. R. & FREIER, N. G. 2006. Robotic pets in the lives of preschool children. *Interaction Studies,* 7**,** pp. 405-436.

KATZ, B. 2004. Development and Current Issues of Interactive Television in the UK. *In:* MASTHOFF, J., GRIFFITHS, R. & PEMBERTON, L., eds. European interactive television conference: Enhancing the experience,Brighton

KING, J. & ALLOWAY, N. 1992. Preschoolers Use of Microcomputers and Input Devices. *Journal of Educational Computing Research,* 8**,** pp. 451-468.

KLOO, D., PERNER, J., KERSCHHUBER, A., DABERNIG, S. & AICHHORN, M. 2008. Sorting between dimensions: Conditions of cognitive flexibility in preschoolers. *Journal of Experimental Child Psychology,* 100**,** pp. 115-134.

KNELL, G. E. 2006. The New Frontier of Educational Television. *Children Now,* 1**,** pp. 1-8.

KUNERT, T. 2009. *User-Centred Interaction Design Patterns for Interactive Digital Television Applications,* London, Springer.

KUNERT, T. & KRÖMKER, H. 2008. Proven Interaction Design Solutions for Accessing and Viewing Interactive TV Content Items. *In:* LEKAKOS, G., CHRORIANOPOULOS, K. & DOUKIDIS, G. (eds.) *Interactive Digital Television: Technologies and Applications.* London: IGI Publishing.

LAKOFF, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind,* London, The University of Chicago Press.

LARGE, A. & BEHESHTI, J. 2005. Interface Design, Web Portals, and Children. *Library Trends* 54 pp. 318–342.

LIVINGSTONE, S. 2004. *UKCGO (UK Children Go Online) Parent questionnaire* [Online]. London: LSE Research Online. Available: http://eprints.lse.ac.uk/502/1/ParentQuestionnaire.pdf [Accessed 16/07 2009].

LIVINGSTONE, S. & BOVILL, M. 2000. Young People, New Media. London: Media Culture.

LIVINGSTONE, S. & HADDON, L. 2009. EU Kids Online: Final report. London: LSE.

LU, C. & FRYE, D. 1992. Mastering the Machine: A Comparison of the Mouse and Touch Screen for Children's Use of Computers. *ICCAL, the International Conference on Computers and Learning.* Wolfville, Canada: Springer-Verlag.

LU, K. Y. 2005. *Interaction Design Principles for Interactive Television* Master of Science in Information Design and Technology, Georgia Institute of Technology.

LUCKIN, R., BOULAY, B. D., YUILL, N., KERAWALLA, C., PEARCE, D. & HARRIS, A. 2003a. Using Software Scaffolding to Increase Metacognitive Skills amongst Young Learners *AIED Artificial Intelligence in Education.* Sydney, Australia.

LUCKIN, R., CONNOLLY, D., PLOWMAN, L. & AIREY, S. 2002. The Young Ones: The implications of media convergence for mobile learning with infants. *In:* European Workshop on Mobile and Contextual Learning.

LUCKIN, R., CONNOLLY, D., PLOWMAN, L. & AIREY, S. 2003b. Children's interactions with interactive toy technology. *Journal of Computer Assisted Learning,* 19**,** pp. 165-176.

LURY, K. 2002. Children's Channels. *In:* BUCKINGHAM, D. (ed.) *Small Screens: Television for Children.* London: Leicester University Press.

LYMAN, M. & LEWANDOWSKI, G. 2005. Genetic programming for association rules on card sorting data. *In:* Proceedings of the 2005 conference on Genetic and evolutionary computation,Washington DC, USA. ACM, pp. 1551-1552.

MACCOBY, E. E. 1951. Television: Its Impact on School Children. *American Association for Public Opinion Research,* 15**,** pp. 421-444.

MALONE, T. W. 1982. Heuristics for designing enjoyable user interfaces: Lessons from computer games. *In:* Proceedings of the 1982 conference on Human factors in computing systems,Gaithersburg, Maryland, United States. ACM Press, pp. 63-68.

MANSOR, E. I., ANGELI, A. D. & BRUIJN, O. D. 2008. Little fingers on the tabletop: A usability evaluation in the kindergarten. *In:* Horizontal Interactive Human Computer Systems, 2008. TABLETOP 2008. 3rd IEEE International Workshop,Amsterdam.

MARKOPOULOS, P., READ, J. C., MACFARLANE, S. & HÖYSNIEMI, J. 2008. *Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers,* Burlington, MA, Morgan Kaufmann.

MASTHOFF, J. & PEMBERTON, L. 2005. Adaptive Hypermedia for Personalized TV. *In:* CHEN, S. Y. & MAGOULAS, G. D. (eds.) *Adaptable and Adaptive Hypermedia Systems.* London: IDEA group

MCKENNEY, S., NIEVEEN, N. & AKKER, J. V. D. 2006. Design Research from a Curriculum Perspective. *In:* AKKER, J. V. D., GRAVEMEIJER, K., MCKENNEY, S. & NIEVEEN, N. (eds.) *Educational Design Research.* Oxon: Routladge

MEGGITT, C. 2007. *Child Development,* Oxford, Heinemann.

MERVIS, C. B. & PANI, J. R. 1980. Acquisition of Basic Object Categories. *Cognitive Psychology,* 12**,** pp. 496-522.

MURRAY, G. C. & REUTER, K. 2005. Children's acquisition of categories and the implications for research in the development of classification schemes. *Annual Meeting of the American Society for Information Science & Technology.*

NGUYEN, S. P. & MURPHY, G. L. 2003. An Apple is More Than Just a fruit: Cross-Classification in Children's Concepts. *Child Development,* 74**,** pp. 1783-1806.

NIELSEN, J. 1995. *Card Sorting to Discover the Users' Model of the Information Space* [Online]. Available: http://www.useit.com/papers/sun/cardsort.html [Accessed 05th May 2008].

NIELSEN, J. 2005. *Ten Usability Heuristics* [Online]. Available: http://www.useit.com/papers/heuristic/heuristic_list.html [Accessed 2 July 2007].

NIELSEN, J. & MACK, R. L. 1994. *Usability Inspection Methods,* New York, NY, John Wiley & Sons.

NIELSEN, J. & MOLICH, R. 1990. Heuristic evaluation of user interfaces. *In:* Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people,Seattle, Washington, United States. ACM Press, pp. 249-256.

NIJHOLT, A. & TAN, D. 2008. Brain-Computer Interfacing for Intelligent Systems. *IEEE Intelligent Systems,* 23**,** pp. 72-79.

NOBLE, G. 1975. *Children in Front of the Small Screen,* London, Constable.

NORMAN, D. 1983. Some Observations on Mental Models. *In:* GENTNER, D. & STEVENS, A. L. (eds.) *Mental Models.* Hillsdale, NJ: Lawrence Erlbaum.

OFCOM 2009. The Communications Market: Digital Progress Report Digital TV, Q1 2009. Office of Communications.

OSWELL, D. 2002. *Television, Childhood and the Home. A History of the Making of the Child Television Audience in Britain,* New York, Oxford University Press.

PAPERT, S. 1980. *Mindstorms: children, computers, and powerful ideas,* Basic Books, Inc.

PECORA, N., MURRAY, J. P. & WARTELLA, E. A. 2007. *Children and Television: Fifty Years of Research,* Mahwah, New Jersey, Lawrence Erlbaum.

PEREIRA, R. C. 2002. A Erotização Precoce e o Papel da Televisão Brasileira. Rio de Janeiro: Universidade Candido Mendes.

PLOMP, T. 2007. Educational Design Research: an Introduction. *In:* PLOMP, T. & NIEVEEN, N., eds.,Shanghai, China. SLO Netherlands institute for curriculum development.

PLOWMAN, L. 1991. *An Investigation of Design Issues for Group Use of Interactive Video.* Doctor of Philosophy, Brighton Polytechnic.

PLOWMAN, L., LUCKIN, R., LAURILLARD, D., STRATFOLD, M. & TAYLOR, J. 1999. Designing multimedia for learning: narrative guidance and narrative construction. *In:* Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit,Pittsburgh, Pennsylvania, United States. ACM, pp. 310-317.

PLOWMAN, L. & STEPHEN, C. 2003. A 'benign addition'? Research on ICT and pre-school children. *Journal of Computer Assisted Learning,* 19**,** pp. 149-164.

PLOWMAN, L. & STEPHEN, C. 2005. Children, play, and computers in pre-school education. *British Journal of Educational Technology,* 36**,** pp. 145-157.

PREECE, J., ROGERS, Y. & SHARP, H. 2002. *Interaction design: beyond human-computer interaction,* New York, John Wiley.

READ, J. C. 2005. *Handwrit ing Recognition Technology, Children, and the Writ ing Process.* Doctor of Philosophy, University of Central Lancashire.

READ, J. C. & MACFARLANE, S. 2006. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. *In:* Proceeding of the 2006 conference on Interaction design and children,Tampere, Finland. ACM Press, pp. 81-88.

REES, D. 2001. *Essential Statistics,* London, Chapman and Hall.

REEVES, T. C. 2006. Design research from a technology perspective. *In:* AKKER, J. V. D., GRAVEMEIJER, K., MCKENNEY, S. & NIEVEEN, N. (eds.) *Educational Design Research.* Oxon: Routledge.

REISER, R. A., TESSMER, M. A. & PHELPS, P. C. 1984. Adult-Child Interaction in Children's Learning From "Sesame Street". *ECTJ,* 32**,** pp. 217-223.

REVELLE, G. L. 2003. Educating via entertainment media: the Sesame Workshop approach. *ACM Computers in Entertainment,* 1**,** pp. 1-9.

REVELLE, G. L. & STROMMEN, E. F. 1990. The Effects of Practice and Input Device Used on Young Children's Computer Control. *Journal of Computing in Childhood Education,* 2**,** pp. 33-41.

RICE, M. & ALM, N. 2007. Sociable TV: Exploring User-Led Interaction Design for Older Adults. *In:* CESAR, P., CHRORIANOPOULOS, K. & JENSEN, J. F., eds. Euro iTV,Amsterdam, The Netherlands. Springer, pp. 126-135.

RICE, M. L., HUSTON, A. C., TRUGLIO, R. & WRIGHT, J. 1990. Words from "Sesame Street": Learning Vocabulary While Viewing. *Developmental Psychology,* 26**,** pp. 421-28.

RIDEOUT, V. & HAMEL, E. 2006. The media family: Electronic media in the lives of infants, toddlers, preschoolers and their parents. USA: Kaiser Family Foundation.

RIEBER, L. P. 1996. Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research & Development,* 44**,** pp. 43-58.

ROBINSON, T. 2001. Television Viewing and Childhood Obesity *Pediatric Clinics of North America* 48**,** pp. 1017 - 1025.

ROMEO, G., EDWARDS, S., MCNAMARA, S., WALKER, I. & ZIGURAS, C. 2003. Touching the screen: issues related to the use of touchscreen technology in early childhood education. *British Journal of Educational Technology,* 34.

RYOKAI, K., VAUCELLE, C. & CASSELL, J. 2003. Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning,* 19**,** pp. 195 - 208.

SCAIFE, M. & ROGERS, Y. 1999. Kids as Informants. *In:* DRUIN, A. (ed.) *The Design of Children's Technology.* San Francisco: Morgan Kaufmann Publishers.

SCAIFE, M., ROGERS, Y., ALDRICH, F. & DAVIES, M. 1997. Designing for or designing with? Informant design for interactive learning environments. *In:* Proceedings of the SIGCHI conference on Human factors in computing systems,Atlanta, Georgia, United States. ACM Press, pp. 343-350.

SCHAFFER, H. R. 2004. *Introducing Child Psychology,* Oxford, Blackwell Publishing

SHAFFER, D. R. 1999. *Developmental Psychology: Childhood and Adolescence* Pacific Grove, CA, Brooks/Cole Publishing

SHERIDAN, M. D., FROST, M. & SHARMA, A. 1997. *From Birth to Five Years: Children's Development Progress,* London, Routledge.

SHIN, S., YIM, J. & SONG, S. 2008. Mu: channel UI to optimize the widget control in internet tv. *In:* Proceeding of the 1st international conference on Designing interactive user experiences for TV and video,Silicon Valley, California, USA. ACM, pp. 41-44.

SHNEIDERMAN, B. 1998. *Designing the User Interface: Strategies for Effective Human-Computer Interaction,* Reading, Massachusetts, Addison Wesley.

SIGMAN, A. 2007. Visual voodoo: the biological impact of watching TV. *Biologist* 54**,** pp. 12-18.

SILVER, M. 2005. *Exploring Interface Design: Proven Techniques for Creating Compelling & Usable Interfaces for Multimedia & the Web,* New York, Thomson Delmar Learning.

SMITH, L. 2005. Emerging Ideas about Categories *In:* GERSHKOFF-STOWE, L. & RAKISON, D. H. (eds.) *Building Object Categories in Developmental Time.* London: Lawrence Erlbaum Associates.

SMITH, P. K., COWIE, H. & BLADES, M. 2003. *Understanding children's development,* Oxford, Blackwell.

SOMERVILLE, I. 1995. *Software Engineering,* New York, NY, Addison-Wesley.

SPRENT, P. 1990. *Applied Nonparametric Statistical Methods,* London, Chapman and Hall.

STATISTICS, N. 2009. Internet Access: Households and Individuals. Newport: Office for National Statistics.

STEIN, S. B. 1979. *Learn at Home the Sesame Street Way,* New York, Simon and Schhuster.

STEINER, K. E. & MOHER, T. G. 1992. Graphic StoryWriter: an interactive environment for emergent storytelling. *In:* Proceedings of the SIGCHI conference on Human factors in computing systems,Monterey, California, United States. ACM Press, pp. 357-364.

STROMMEN, E. 1998. When the interface is a talking dinosaur: learning across media with ActiMates Barney. *In:* Proceedings of the SIGCHI conference on Human factors in computing systems,Los Angeles, California, United States. ACM Press/Addison-Wesley Publishing Co., pp. 288-295.

SUNG, Y.-T., CHANG, K.-E. & LEE, M.-D. 2008. Designing multimedia games for young children's taxonomic concept development. *Computers & Education* [Online], 50.

TUCHSCHERER, P. 1988. *TV Interactive Toys: The New High Tech Threat to Children,* Bend, Oregon, Pinnaroo.

VYGOTSKY, L. S. 1978. *Mind in Society: The Development of Higher Psychological Processes,* Cambridge, Harvard University Press.

WEERAMANTHRI, N. 2008. The rise and Fall of Juan Perez: Key Pressing Behaviour and Children's Responses to Interactive Television. *In:* TSCHELIGI, M., OBRIST, M. & LUGMAYR, A., eds. Euro iTV,Salzburg, Austria. Springer Berlin, pp. 20-29.

WHARTON, C., RIEMAN, J., LEWIS, C. & POLSON, P. 1994 The cognitive walkthrough method: a practitioner's guide *In:* NIELSEN, J. & MACK, R. L. (eds.) *Usability inspection methods.* New York: John Wiley & Sons, Inc.

WILSON, K. S. 1988. Palenque: an interactive multimedia digital video interactive prototype for children. *In:* Proceedings of the SIGCHI conference on Human factors in computing systems,Washington, D.C., United States. ACM, pp. 275-279.

WRIGHT, J. C., HUSTON, A. C., MURPHY, K. C., PETERS, M. S., PIÃ±ON, M., SCANTLIN, R. & KOTLER, J. 2003. The Relations of Early Television Viewing to School Readiness and Vocabulary of Children from Low-Income Families: The Early Window Project. *Child Development,* 72**,** pp. 1347-1366.

WYETH, P. & WYETH, G. 2001. Electronic Blocks: Tangible Programming Elements for Preschoolers. *In:* Eighth IFIP TC.13 Conference On Human-Computer Interaction (INTERACT 2001) Tokyo. pp. 496-503.

XU, J. A. & ARAKI, K. 2005. A personalized recommendation system for electronic program guide. *In:* ZHANG, S. & JARVIS, R., eds. 18th Australian Joint Conference on Artificial Intelligence,Sydney, AUSTRALIA. Springer-Verlag Berlin, pp. 1146-1149.

ZIMMERMAN, F. J. & CHRISTAKIS, D. A. 2005. Children's Television Viewing and Cognitive Outcomes A Longitudinal Analysis of National Data. *ARCH PEDIATR ADOLESC MED* 159**,** pp. 619-626.