COMPARING CHILDREN'S ENJOYMENT AND ENGAGEMENT USING PHYSICAL, GRAPHICAL AND TANGIBLE USER INTERFACES

by

Zhen (Lesley) Xie
B. Eng., Shanghai Jiaotong University 2000

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
School of Interactive Arts and Technology

SIMON FRASER UNIVERSITY

Spring 2008

APPROVAL

Name: **Zhen (Lesley) Xie**

Degree: **Master of Science**

Title of Thesis: **Comparing Children's Enjoyment and Engagement Using Physical, Graphical and Tangible User Interfaces**

Examining Committee:

____

Chair: Ron Wakkary
Associate Professor, School of Interactive Arts and Technology

____

Dr. Alissa Antle
Senior Supervisor
Assistant Professor, School of Interactive Arts and Technology

____

Dr. Chris Shaw
Supervisor
Associate Professor, School of Interactive Arts and Technology

____

Dr. Alyssa Wise
External Examiner
Assistant Professor, School of Education

Date Defended/Approved: ______________________________

ABSTRACT

This paper presents the results of an exploratory comparative study in which I investigated the relationship between interface style and school-aged children's enjoyment and engagement while doing puzzles. Pairs of participants played with a jigsaw puzzle which was implemented using three different interface styles: physical (traditional), graphical and tangible. In order to investigate interactional differences between the three interface styles, I recorded subjective ratings of enjoyment and three related subscales, and measured times and counts of engagement. Qualitative analysis based on observational notes and audio responses to open interview questions helped contextualize the quantitative findings and provided key insights into interactional differences not apparent in the quantitative findings. I summarized the main findings and discussed the design implications for tangible user interfaces. The main contribution of the study documented in this paper is that it is the first empirical comparison of physical (traditional), graphical and tangible interfaces for school-aged children. A second contribution is the development of an extensible tabletop prototype, which uses fiducial markers and a camera vision system to track user driven events. The third contribution is the set of design recommendations for the development of enjoyable and engaging tangibles.

Keywords: Interface style, enjoyment, engagement, collaboration, gender, children, play, puzzles, tangible user interfaces.

Subject Terms: Interface style – tangible user interfaces (TUIs), graphical user interfaces (GUIs), physical (traditional) user interfaces (PUIs), Enjoyment – interest and enjoyment, perceived competence, perceived choice, pressure and tension, Engagement, Collaboration, Gender effect.

ACKNOWLEDGEMENTS

This research study was funded by the National Science and Engineering Research Council of Canada (NSERC). I am deeply thankful to all those who have strongly supported and encouraged me through my two-year graduate studies. Special thanks to Alissa Antle, my senior supervisor, for her much appreciated expert mentorship and support of my work step by step for all these years. Many thanks to Nima Motamedi, who provided his help to develop a novel tabletop system that is used in this research study. Deep thanks to Caitlin Akai, who helped to facilitate all the user studies run at Science World at the Telus World of Science.

Outside the school, I hold deep gratitude to Sandy Eix, Rob Lunde and all the staff working at Science World during our user study time. They facilitated us successfully running the user studies. I also thank all the children and their parents who voluntarily participated to my research studies.

On a personal note, I'd like to acknowledge the friendship and support of many school friends who have made my work and time at SIAT a really great experience. Some of them also provided valuable ideas to my research study. They are: Emily Yao, Huaxin Wei, Milena Droumeva, Greg Corness, and countless others.

Last but not least, special thanks give to my husband, Jack Zhang for always being there for me, supporting all my lives; to my parents and my parents-in-law, who strongly supported my studies during my whole study period.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Tangible User Interface	Utilizing physical representation and manipulation of digital data, offering interactive couplings of physical artifacts with "computationally mediated digital information".
Enjoyment	Enjoyment relates to intrinsic motivation, which describes natural inclinations toward spontaneous interest and exploration that are essential to cognitive and social development.
Engagement	Engagement comprises cognitive engagement, which involves attention to the activity and concentration and promotes 'useful' learning.

CHAPTER 1: INTRODUCTION

1.1 Overview

Computation has been used to augment children's play in a variety of ways (Fails et al., 2005; Montemayor et al., 2002; Papert, 1980). A recent trend is the application of tangible user interfaces (TUIs) to children's learning and play-based applications and products (e.g. Antle, 2007; Marshall, 2007; Rogers et al., 2006; Zuckerman et al., 2005; Resnick et al., 1998). The work presented in this thesis was primarily driven in response to the emerging trend of designing tangible user interface systems for children. Tangible user interfaces are designed by using physical objects to control and represent digital information. The system allowed users to manipulate this digital information through physical action with these concrete objects.

Much of the current research in this area has focused on the development and description of new tangible user interface systems. Development of tangible user interfaces appear to have been driven by the intuition that tangible style interfaces, which rely on direct physical manipulation and support face-to-face collaboration, are more "natural" or intuitive and thus more beneficial to enjoyment and engagement on playful learning tasks for children than desktop environments. However, it is unclear whether this general assumption is correct. Few empirical studies have addressed these claims (Antle, 2007; Marshall, 2007). Compared to graphical style desktop systems, there has been little empirical research that explicitly and systematically explores the advantages of tangible user interfaces. The claims of the benefits of tangible interaction remain speculative. It is unknown how the properties of tangible interaction will contribute to enjoyment and engagement in tangible games for school age children. Understanding these issues is important because it will contribute to grounding this technology agenda in empirical studies; inform the development of stronger frameworks for the theory and practice of play-based learning with tangibles; and lead to the development of principles to guide the design of new forms of tangibles. The results of empirical studies can provide guidelines which should be considered in the design of the tangible games for children in the future. All these issues outlined above form the foundation for my research.

Although there have been a few theoretical claims of tangible user interfaces implicitly proposed in the last decade, little evidence gathered through empirical work is found to date. In the absence of validated theoretical frameworks, it is difficult for the research designer to determine whether choosing to develop a tangible user interface for children will bring substantial benefits to the young user for playful learning tasks. If a tangible user interface design is chosen, what characteristics of tangible user interfaces might contribute to playful learning, and which factors of tangible user interfaces will provide most benefit?

Enjoyment and engagement are integral and prerequisite aspects of children's playful learning experiences (Malone, 1980; Prensky, 2001; Heidegger, 1990; Montessori, 1965). They both are required for informal learning and may be a benefit to actual learning as well. Thus, I chose them as two primary dependent variables evaluated in this research study.

I chose spatial games because a tangible user interface is a logical choice for spatial activities due to its direct access to physical manipulation and support of face-to-face collaboration especially for child users. Jigsaw puzzles were chosen as they represent a familiar playful activity that is undertaken socially, requires cognitive effort, utilizes physical manipulation, and is spatial in nature.

This thesis describes a controlled comparative study exploring how interface style related interaction factors impact enjoyment and engagement in jigsaw puzzle games under a collaborative condition for 132 school aged children (7-9 years). The main contribution of this thesis is that it is the first empirical comparison of physical (traditional), graphical and tangible interfaces for school-aged children. A second contribution is the development of an extensible tabletop prototype, which uses fiducial markers and a camera vision system to track user driven events, such as the connection of two or more puzzle pieces. A third contribution is a set of design recommendations for the development of enjoyable and engaging tangibles.

1.2 Research Goal

The overarching goal of this research is to explore and investigate the relation between interface styles and school-aged children's enjoyment and engagement on a playful learning task. The design of this study attempts to advance the understanding of how tangible user interfaces could be used for playful learning tasks when designed for child users.

The main research question I explore in this study is:

Does interface style (i.e., physical (traditional) user interface, graphical user interface and tangible user interface) affect children's enjoyment and engagement on playful learning tasks?

1.3 Thesis Guide

A brief overview of how this thesis is structured is given below.

In Chapter Two, the background chapter, I outline literature in several areas that are related to the aims of this research study. First, I present theories and previous research in the field of tangible user interfaces, tabletop systems and tangible interaction design for children. Research on interaction design and evaluation for children are then reviewed. The review of these studies provides design guidelines and validated evaluation approaches that can be used for child users. In the next section of this chapter, I summarize the theoretical and operational definitions of enjoyment and engagement, which are the two main dependent variables analyzed in this study. I then look at literature based on children's collaboration in learning. I outline the definitions of two important collaboration styles used in the study, which are independent parallel play and sequential turn-taking. I also present a review of literature related to gender. In the final section of this chapter, I state the motivation of this research study and conclude with a main research question followed by six specific questions related to this study.

In Chapter Three, I outline the experimental framework used to conduct my research. I begin by presenting five main hypotheses I formulated to address my research questions. I then introduce the design of jigsaw puzzles implemented using three different interface styles. I focus on the features and characteristics that are relevant to my experiment. In the remainder of the chapter, I describe the sets of tasks and the measures used for evaluating children's enjoyment, engagement and collaboration.

In Chapter Four, I extend the framework given in Chapter Three by describing, in detail, the experimental methodology used to examine the relation between interface style and subjects' enjoyment, engagement and collaboration. I provide detailed information on the experimental setup, including a description of test subjects, test environment, lab setup, assumptions, the pilot study and the design of subsequent main experiment. I also provide a detailed discussion of the statistical analysis tools that are used for further examination.

In Chapter Five, I present the quantitative statistical analysis of the two main variables of enjoyment and engagement. I conclude with the analysis results from each part of the experiment. I also explore the qualitative findings based on observational notes and audio recordings to generate the findings on children's enjoyment, engagement and collaboration.

In Chapter Six, I provide interpretations and comments on all the findings and evidence from this study in the context of my original research questions. I discuss the implications of these results related to the design of tangible user interfaces. I compare the findings with previous related research studies in the field of interaction design for children.

In the final chapter, Chapter Seven, I conclude the findings and summaries of contributions. I then state suggestions for future studies in this research field. I also indicate the limitations of this study and present some fresh ideas for future explorations.

In the Appendices section, I provide the session scripts and the lists of pre- and post questionnaires which were used during the experimental study.

CHAPTER 2: THEORETICAL BACKGROUND

2.1 Overview

The design idea of this study was inspired by the lack of empirical experiments exploring the benefits of building tangible user interfaces for young users. In this chapter, I present the background theories and previous studies in the fields of tangible user interfaces research for children. Firstly, I introduce the field of tangible user interfaces and summarize research to date. Secondly, I outline recent trends in interaction design for children and discuss the successes and pitfalls of research studies in this field. Lastly, I describe the application of these guidelines to the design of this research and pose specific research questions. The design of the experimental framework for this study was informed by the design guidelines derived from the literature I discuss in this chapter.

In Section 2.2 of this chapter, I introduce the field of tangible user interfaces and summarize research to data. In Section 2.3, I discuss the growing interest in interactive tabletops with their support for co-located sharable activities. In Section 2.4, I describe the recent trend of developing tangible user interfaces for children and discuss the pros and cons of several noteworthy empirical studies in this field. In Section 2.5, I outline some recent studies in interaction design and validated evaluation methods for children. In Section 2.6, I introduce the two primary dependent variables – enjoyment and engagement, and outline the conceptual and operational definitions of these two variables. In Section 2.7, I explore children's collaboration in informal learning with interactive technologies. Such collaboration is predicted to be a benefit of tangible user interfaces and is frequently discussed in this research field. I predict that two important collaboration styles, which are parallel independent play and sequential turn-taking, will be commonly observed in the study. I also discuss the conceptual definitions of these two styles given by other researchers. In Section 2.8, I review some empirical studies related to gender. Gender is often discussed in interaction design studies for children. This is because many researchers find gender disparities exist with respect to interactions with computers, especially for children (Cassell et al., 1998). In the remainder of the chapter, I continue to address the research motivation (Section 2.9), describe the implementation of design guidelines to the design of this research study and outline all the research questions (Section 2.10) of this study.

In the following chapter, Chapter Three, I describe in detail the design framework of this study, which is used to examine all the research questions I introduce in this chapter.

2.2 What are Tangible User Interfaces (TUIs)?

The last two decades have seen a wave of new research dealing with the coupling the physical and digital worlds. The terms defined in early research were varied, such as "passive real-world props" (Hinckley et al., 1994), "graspable user interfaces" (Fitzmaurice et al., 1995), "manipulative user interfaces" (Hinckley et al., 1994), and "embodied user interfaces" (Fishkin et al., 1998). However, these multiple terms are largely synonymous without any differences. They all share the same basic paradigm in which a user uses his/her hands to manipulate some physical object(s) through physical gestures, and a computing system detects the manipulation, alters its state and gives feedback accordingly (Fishkin, 2004). These types of interfaces allow users to take advantage of their spatial skills and to interact collaboratively with augmented physical objects in order to access and manipulate digital information. Consolidating upon this foundation, Ullmer and Ishii (1997) proposed the term

"Tangible User Interfaces (TUIs)". They and others (Hornecker et al., 2006; Holmquist et al., 1999) defined TUIs as interfaces that utilize physical representation and manipulation of digital data, and offer interactive couplings of physical artifacts with "computationally mediated digital information" (Holmquist et al., 1999). Since then, the terms of *Tangible User Interfaces* and *Tangible Interaction* have increasingly gaining currency within the Human Computer Interaction (HCI) domain (Hornecker et al., 2006).

Traditional graphical user interfaces (GUIs) define a set of graphical interface elements (e.g., windows, icons, menus) that reside in a purely electronic or virtual form. Generic indirect input devices (e.g., mouse and keyboard) are primarily used to manipulate these virtual interface elements (Fitzmaurice et al., 1995). The interaction model of GUIs present a strong separation between the digital representation (provided by the graphical display) and control (mediated by the GUI's input devices). Ullmer and Ishii (2001) illustrated the relationship between its "input" and "output" components by the "model-view-control" archetype (Figure 2.1).

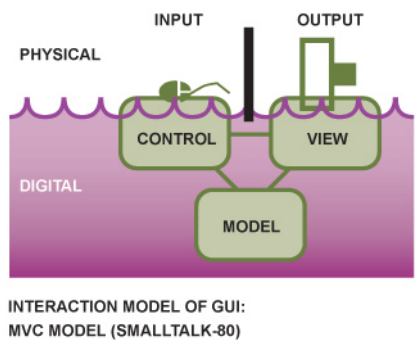

Figure 2.1 Graphical User Interfaces (GUIs) interaction model (Ullmer and Ishii, 2001)

Ullmer and Ishii (2001) presented a narrower definition of tangible user interfaces and proposed a conceptual framework for this paradigm. They defined tangible user interfaces as a user interface that couples physical representations (i.e., spatially manipulable physical objects) with digital representations (i.e., graphics and audio), yielding interactive systems that are computationally mediated, but generally not identifiable as "computers" per se. In terms of the external representation, they divided it into two broad classes, which are physical representation and digital representation. They consider physical representation to be information that is physically embodied in concrete, "tangible" form, and digital representation to be computationally mediated displays that are perceptually observed in the world, but are not physically embodied, and thus "intangible" in form (Ullmer et al, 2001).

The interaction model for tangible user interfaces Ullmer and Ishii (2001) presented depicted the integration of physical representation and control. It is the most cited of conceptual frameworks in recent HCI publications. The viewpoint of this conceptual framework is from a data-centred view in which they explored the possibility of coupling between material and virtual representations. Figure 2.2 represents the interaction model for tangible interfaces by a "model-control-representation (physical and digital)" archetype (Ullmer and Ishii, 2001).

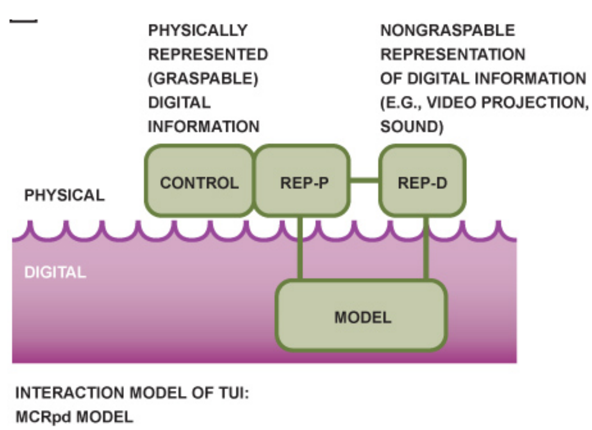

Figure 2.2 Tangible User Interfaces (TUIs) interaction model (Ullmer and Ishii, 2001)

Broader characterizations of tangible interfaces have been instantiated in design frameworks, which concentrate on the design of the interaction itself (Antle, 2007). Fishkin (2004) presented a taxonomy, which used metaphor and embodiment as its two axes and unified a series of previous categorizations and definitions. The novel taxonomy unifies several previous frameworks, naturally lends itself to design principles to guide future studies in the field, reveals new structure in the history of the work in the field, and seamlessly integrates "more tangible" interfaces with "less tangible" ones. Hornecker et al. (2006) provided a framework that encompassed a broad range of mixed reality interfaces and systems relying on embodied interaction, tangible manipulation and physical representation (of data), embeddedness in real space and digitally augmenting physical spaces. They discussed the role of collaborative and spatial aspects, and provided a good overview of previous approaches to conceptualizing tangible and spatial interaction.

2.3 Interactive Tabletops

Recently, there has been a growing interest in how large shared displays can be designed to facilitate small co-located group working, including the use of tabletop displays. Compared with traditional personal computers, interactive tabletops have been found to encourage contributions from all group members and to support more equitable problem solving and decision-making (e.g., Hornecker, 2005; Khandelwal, 2006). Research concerned with extending the tabletop has focused on how to integrate it with other devices, including personal computers, handhelds, tangibles and augmented reality (e.g., Hornecker, 2005; Sugimoto et al., 2004; Ishii et al., 1997).

A popular paradigm for tangible interfaces is based upon the concept of "interactive surfaces", where physical objects are manipulated by users on an augmented surface. A number of tangible interfaces have also been built based on a tabletop surface. Many different research projects, including Ishii and Ulmer (1997), FitzMaurice et al. (1995), Rauterberg et al. (1998), Aliakseyeu et al. (2001), and Mazalek et al. (2003), have studied the required technology, usability and possible applications of tangible tabletop interfaces. Projects utilizing augmented tabletop environments have demonstrated their potential value over the past ten years. Rogers et al. (2006) discussed the benefits of extending the tabletop into a physical-digital space. They conclude that a major advantage of tabletop environments is that it opened up more opportunities for synchronous collaborative tasks, inviting all to browse, pick up, pass around and compare options. However, little is known about how and why such environments can be designed to support successful social interactions. Terrenghi et al. (2007) presented a comparative study in which they investigated the ways manipulation of physical versus digital media are fundamentally different from one another. They observed that bimanual interaction in a digital domain was largely symmetric in nature, which was quite different from the kind of asymmetric bimanual interaction typical of physical manipulation. They also found that users predominantly used one hand in 2-dimensional digital interaction, but used two hands together both in 3-dimensional space and with tactile objects. They suggested that physical metaphors and methods of input might appear to encourage manipulation in a physical way, but in the digital realm it was essentially quite different.

Scott et al. (2003) presented a critical analysis of the current state-of-the-art in digital tabletop systems research, targeted at discovering how user requirements for collaboration are currently being met and uncovering areas requiring further development. By considering research on tabletop displays, collaboration, and communication, they proposed eight guidelines for effective co-located collaboration around a tabletop display. These design guidelines stress the importance of allowing users to easily integrate the collaboration with their tasks and supporting users' familiar work practices at the tables. For example, users can use physical objects and sit at different positions around the table. These guidelines are partially in line with some characterizations of tangible interfaces. A tabletop system was considered to be a logical solution for this study.

2.4 Tangible User Interfaces for Children

Many people have explored how technology can enhance learning during children's play, the role technology can and should play, and how best to support children to develop cognitively through augmented play activities (Fails et al., 2005). Healy (1998) provided her support for this research area when she stated that body movements, the ability to touch, feel, manipulate and build sensory awareness of the relationships in the world was crucial to children's cognitive development (Antle, 2007).

Building TUI systems specifically for children has become popular in the last decade. Tangible systems are considered to have a powerful ability to engage school age children in active play, which promotes cognitive development. The development of tangible systems specifically targeted to children is also a growing research area. This research builds on past research themes, which have explored how technology can enhance learning during children's play; the role technology can and should play in children's lives; and how children can be supported to develop cognitively through augmented play activities.

Most research on tangible and spatial user interaction for children focuses on building systems rather than proposing explanations for how and why tangibles might cause particular learning effects (Antle, 2007). There has been a rise of research based on developing a conceptual and theoretical understanding of designing tangible interaction specifically for children in these years. Zuckerman et al. (2005) introduced a new framework for thinking about tangible interfaces in education, with a specific focus on abstract problem domains. They presented the classification of tangible manipulatives as "Froebel-inspired Manipulatives" (FiMs) or "Montessori-inspired Manipulatives" (MiMs) depending on intended use. They argued that FiMs are design materials, fostering modelling of real-world structures, while MiMs foster modelling of more abstract structures. Rogers et al. (2002) presented a conceptual framework for mixed reality specifically for children. It focuses on the notion of transformations between virtual and physical dimensions. Marshall (2007) suggested the possibility of using the Heidegger's distinction between "readiness-to-hand" and "presence-at-hand" to promote reflection in children. What is missing in these previous studies is a design framework grounded in child-specific developmental theories about how children develop intelligence through their physical, social and spatial interaction with the world. Antle (2007) provided another conceptual framework for the design of tangibles and interactive spaces, which support schemata level knowledge acquisition in children. The CTI framework is presented in five themes. Each of the five themes relates to a feature or aspect of tangible systems. These five dimensions of the conceptual framework defined vertical guidelines for tangible and spatial interaction and children and informed the design of this research study. Four of the important features, including direct physical manipulation, integration of input and output space, face-to-face collaboration, and digital feedback, were used for the design of the tangible interface that was tested in this research study. More details are discussed in section 3.4.

Some of the earliest studies attempted to explore the potential benefits of tangible user interfaces for supporting children's play and learning. For example, Price et al. (2003) designed an adventure game on a TUI for children in order to understand the effectiveness of tangible-mediated learning, in terms of its particular benefits. They claimed that interaction with tangibles encouraged engagement, excitement and collaboration. They also suggested that tangibles can be used to create novel learning environments, which have the potential to make learning playful and pleasurable through engaging children in exploratory and reflective activities. However, all measures were qualitative and subjective. Africano et al. (2004) implemented and evaluated a design of tangible interfaces as a learning tool for supporting school aged children to collaboratively learn on school tasks. She concluded that tangible systems can support children's collaboration, enjoyment and engagement in the play, and thus the system can make a positive contribute to the existing array of learning tools. However, Africano et al.'s (2004) operationalization of engagement was subjective. Furthermore, none of these previous studies explores how and why tangibles can support playful learning activities for children.

Many different research projects have studied enabling technologies, usability aspects and various applications of tangible user interfaces especially for children. There are many noteworthy studies. For example, McNerney (2004) provided a historical overview of educational computing research at MIT from the mid-1960s to the present day, focusing on physical interfaces. He suggested that using electronic toys helps children develop advanced modes of thinking through free-form play. He also suggested that, compared to screen-based user interfaces, tangible user interfaces make computation immediate and more accessible, and that they are appropriate for children learning about computation and scientific exploration. His study provides a rich body of research into the field of educational programming and tangible user interfaces. He also provided his viewpoint on the constructionist paradigm in learning through this study. Andersen (2004) observed children's emerging understanding of sensors as they explored and played, and, in the longer term, touchable interfaces. She proposed that objects augmented with electronic sensing capabilities require us to develop new intuitions or "naïve" understandings about both the resulting hybrid object and the specific affordances of the sensor itself. She indicated that touchable interfaces were a useful way to scaffold experience and aid the suspension of disbelief. Bohn (2004) presented a smart jigsaw puzzle assistant that relies on RFID technology to provide a convenient method to interact with the computer. However, evaluations were not performed in the study by using this puzzle. Raffle et al. (2006) conducted a study with children to incorporate robotic creations with modulations of frequency, amplitude, phase and orientation of motion recordings. They suggested that manipulating parameters of motion for mapping enabled children to more deeply design and analyze sophisticated robotic behaviours. They also suggested that making fundamental ideas like phase, amplitude and frequency manipulatable could help older children transfer their knowledge from physical activities to more abstract symbolic representations of movement. Lately, they presented two new composition and performance based tools for robotic control in order to explore users' understanding of robotic manipulation itself through their cooperative and competitive performances. The study concluded that tools providing the means for capturing, organizing and controlling movement in real-time will help children analyze, understand and refine the design of their robotic creations (Raffle et al., 2007). These studies contribute to the growing knowledge base of using tangibles for supporting children's cognitive development on playful learning tasks. However, little is known about the effectiveness of tangible-mediated learning, in terms of its particular benefit over other kinds of interfaces on those learning tasks for children. Fernaeus & Tholander (2006) illustrated a view on tangibles as resources for action instead of only as alternative forms of data representation. Verhaegh et al. (2007) introduced a tangible electronic board game for educational purposes to investigate the balance between challenge and control and its possibility of supporting enjoyable learning.

A number of design-focused projects have suggested that tangible interfaces might be particularly suitable for collaborative learning (e.g., Africano et al., 2004; Fernaeus et al., 2005). Suzuki et al (1995) claimed that tangible interfaces can be designed to create a shared space for collaborative transactions and allow users to monitor each other's gaze to achieve interaction more easily than when interacting with a graphical representation on a display. Other researchers suggested that tangible interfaces might increase the visibility of other member's activity, better communicating the current state of their work and potentially encouraging situated learning (e.g., Fernaeus et al., 2005; Klemmer et al., 2006; Marshall, 2007). Stanon et al. (2001) claimed that collaborative activity might be encouraged by increasing the size of tangible interfaces and using props to slow down the pace of interaction and increase the effort required to make manipulation. However, none of these claims was made based on empirical comparative work. It remains unclear whether tangible interface designs can enhance children's collaborative learning better than any other interface styles (i.e., physical (traditional) user interface or graphical user interface).

Many of these previous studies focus on describing the system and provide descriptive summaries of user interactions rather than proposing explanations for how and why tangibles might cause particular learning effects for children. It was not until recently that researchers start to run some comparative studies with tangible user interfaces over the other traditional user interfaces (e.g. graphical user interface) on specific learning tasks. Fails et al. (2005) presented an exploration of the differences between desktop and physical environments for children. They conducted an experiment in which they compare the use of desktop and physical interactive environments by preschool-aged children. The researchers suggested that embedding technology in the physical world, rather than simply presenting young children with traditional desktop applications, may be beneficial to them. They found that gender could affect the types of interaction with the different environments (physical vs. desktop). They also claimed that comparing desktop and physical environments is difficult and requires multiple metrics. However, so far, there has been few empirical studies to address these claims. The lack of this research propels the design of the research study I present in this paper.

2.5 Evaluation with Children

Children have their own likes, dislikes, curiosities and needs that are not the same as adults (Druin, 2002). Designers of new technologies for young people should be aware that they are not 'just short adults' but an entirely different user population with their own culture, norms and complexities (Berman 1977). In the field of Human Computer Interaction (HCI), growth has occurred in studying children as technology users, which led to efforts to tailor methods for usability testing with children as test participants (MacFarlane et al., 2004).

Druin (2002) suggested a framework for understanding the roles that children can play in the technology design process, particularly in regards to designing technologies that support learning. Markopoulus et al. (2003) continued this study by discussing some perspectives on children's development, their use of technology for entertainment and education, and finally, how to involve children in the various stages of the design progress.

Work by Hanna et al. (1997) produced a set of guidelines for usability testing with children that incorporated general advice on the operation of usability studies. She suggested that children in elementary school age range (ages 6 to 10 years) are relatively easy to include in software usability testing. Hanna et al.'s (1998) later work described a range of techniques for usability research with children (Hanna et al., 1998). These techniques included iterative laboratory tests and longitudinal tests that incorporated questionnaires that were given to the children to complete. Donker & Markopoulos (2001) have investigated the use of thinking aloud, interviewing, and questionnaires as usability evaluation methods for children. These author's recommendations for working with children in experimental setting informed our research design and procedures.

Many researchers have explored the evaluation of fun, play and learning. It is claimed that fun contributes to being motivated to pursue an activity, and as such can also contribute to learning effectively through play for children. Children find informal learning fun when they enjoy and are engaged in the activities. Read, MacFarlane, and Casey (2002) developed a set of tools used in empirical studies for measuring fun with children aged from 5 to 10. They defined fun as three dimensions, which are endurability, engagement, and expectations. Price et al. (2003) described five core elements that can contribute to playful learning, and they are (1) exploration through interaction; (2) engagement; (3) reflection; (4) imagination, creativity, and thinking at different levels of abstraction and (5) collaboration. All these concerns contribute to delimit the scope of my design of this children-related study. We measured "fun" for playful learning tasks with children by using enjoyment and engagement for this research study.

When conducting a research study with children, special evaluation methods are required. This is because child users are different from adults. Researchers developed different kinds of measuring tool kits for children. Risden et al. (1997) proposed a funtool kit, which is called "funometer". It can be used for children to report their satisfaction on play tasks. Read et al. (2002) developed the "smileyometer" based on Risden's "funometer" tool and with the helps of children. The tool is based on a 1-5 Likert scale, and uses pictorial representations of different kind of happy faces to represent the different level of satisfaction. The smileyometer has been proved easy for the children to use and it has been used in different situations to measure one or more of the fun dimensions. Thus, this scale was adopted in this study.

2.6 Enjoyment and Engagement – the two DV

Enjoyment and engagement are two requisites for playful learning activities (Malone, 1980; Prensky, 2001; Heidegger, 1990; and Montessori, 1965). They both are commonly discussed and evaluated in the study of design tangible interaction with children. Fun and enjoyment are well known to be effective in children's development (Clements, 1995), both supporting and deepening learning (Resnick et al., 1999) as well as facilitating engagement and motivation. Dix (2003) discussed the relationship of fun and engagement. He claimed that there are many examples of experiences that were fun and engaging and also experiences that were engaging but not fun; however, it is hard to find things that were fun and not engaging. Prensky (2001) indicated that a combination of twelve elements make games engaging. Fun and enjoyment are the most important elements of all these twelve elements. Being actively engaged in a learning activity has repeatedly been shown to be beneficial for learning (Price et al., 2003). Engagement comprises cognitive engagement, which involves attention to the activity and concentration and promotes 'useful' learning (Stoney et al., 1999). In this research study, enjoyment and engagement are considered to be the two main dependent variables. The conceptual definitions of enjoyment and engagement set the scope and meaning of the terms. Each is a complex construct derived from physical, social and cognitive theories.

2.6.1 Enjoyment

Early research on enjoyment is found in the work of Piaget (1971). He suggested that a child at play repeats his behaviour not in any further effort to learn or investigate but for the mere joy of mastering it. Physical and social scientists have also created a large body of work that relates to enjoyment (e.g., Davis, 1982, Kremer et al., 1997, Wankel, 1993, Ryan, 2000). Davis (1982) presented a causal theory of enjoyment. The basic premise is that an object of enjoyment causes the subject to experience pleasure by causing concurrent beliefs which satisfy desires concerning the experience itself. Pleasure is identified with concurrent happiness, which can be defined in terms of belief, desire, and thought. Degree of enjoyment can be defined as part of the subject's pleasure attributable to the object of enjoyment. In the domain of children's play and learning, an alternative conception is necessary.

Theory and research suggest that the experience of enjoyment is a critical factor in determining one's motivation for and continued participation in exercise settings (Kremer et al., 1997; Wankel, 1993). Weiss (1987) advises that children will be more inclined to participate in physical activity if they perceive it to be enjoyable, and the chance of continued participation will be increased by enhancing their intrinsic motivation. Similar sentiments have been echoed by research in which enjoyment has been found to be positively related to a desire to continue participation (Scanlan et al., 1989), and where the effort to increase intrinsic motivation has been widely accepted as a desirable educational practice since it leads to long-term motivation, and hence continued participation (Deci et al., 1985).

Ryan (2000) developed a model of enjoyment based on self-determination theory, which relates enjoyment to intrinsic motivation. Self-determination theory (SDT) is an approach to human motivation and personality that uses traditional empirical methods while employing an organismic meta-theory that highlights the importance of humans' involved inner resources for personality development and behavioural self-regulation (Ryan, Kuhl, & Deci, 1997). Thus, its arena is the investigation of people's inherent growth tendencies and innate psychological needs that are the basis for their self-motivation and personality integration, as well as for the conditions that foster those positive processes. SDT relates enjoyment (during social activities) with intrinsic motivation. The construct of intrinsic motivation describes natural inclinations toward spontaneous interest and exploration that are essential to cognitive and social development and that represent a principal source of enjoyment (Ryan, 2000).

Intrinsic motivation is a logical choice for evaluating playful learning activities. Ryan (2006) proposed his Intrinsic Motivation Inventory (IMI) tool developed based on the SDT theory. The IMI is a multidimensional measurement device designed for participants' subjective reported experience in laboratory experiments. The instrument includes six subscales, which assess participants' interest and enjoyment, perceived competence, effort, value and usefulness, felt pressure and tension, and perceived choice while performing a given activity. The *Interest and Enjoyment* subscale is considered as the self-reported measure of intrinsic motivation. Moreover, the questionnaire is easy to modify to fit specific activities and interpret for children. Thus, I consider this approach as a useful measure for assessing children's enjoyment of computational games.

Based on the work of Verhaegh et al. (2007) evaluating tangible tabletops for children, we are using four subscales of the Intrinsic Motivation Inventory (IMI) (Ryan, 2006) in this study to evaluate enjoyment. These four subscales provide self reports of participants' *Interest and Enjoyment*, *Perceived Competence*, *Perceived Choice* and *Pressure and Tension* with respect to their experiences playing a game in a research study. *Perceived Competence*, *Perceived Choice* and *Pressure and Tension* subscales are included in the measure, because they are all interrelated to the *Interest and Enjoyment* subscale. Enjoyment in this study was operationally defined as the subject's (children's) ratings of their enjoyment and three relevant subscales on a one to five Likert style scale.

2.6.2 Engagement

There is a growing interest in engagement within the field of human computer interaction. In interaction design studies, engagement has been examined in relation to the usability of digital products (Jacques et al. 1995; Monk, 2002; Rozendaal et al., 2007). However, engagement is still an evasive concept given the multiplicity of meanings that are attributed to it (Hornbæk, 2006; Rozendaal et al., 2007)

Engagement can be described in various ways. Malone (1981) first defined engagement as an exciting and enjoyable state of mind in which attention is willingly given and held. Csikszentmihalyi (1990) addressed engagement as an optimal state of mind, called Flow, in which people report losing the sense of self and time and experiencing effortlessness in the development of skills. For most researchers, "engagement" entails some kind of mindfulness, cognitive effort and deep processing of new information (Salomon et al., 1987). Common in these varied views on engagement is that engaging activities are intrinsically enjoyable, i.e., the activity is performed for intrinsic rewards and is not performed for extrinsic rewards (Deci et al., 2000). Engagement comprises cognitive engagement, which involves attention to the activity and concentration and promotes 'useful' learning (Stoney et al., 1999). This conceptualization is relevant for children's play since a dominant function of play is learning. Learning requires engaged attention.

Some research have operationalized engagement as the amount of time spent on and off a particular task, i.e. actions and discourse not related to the task at hand (Scott et al., 2000). For studies involves children, Hanna et al. (1997) suggested that observed frowns and yawns were more reliable indicators than children's responses to questions. Read et al. (2002) proposed that engagement be measured through a set of positive and negative instantiations, including smiles, laughing, concentration signs, excitable bouncing, and positive vocalization versus frowns, signs of boredom (ear playing, fiddling), shrugs, and negative vocal instantiation. However, when doing study with children, it is hard to determine their real engagement level only by facial expression or behavioural measures. Measures of children's engagement through observations on their behavioural or facial expression are overly subjective and hard to validate.

In order to avoid the biases of subjective measures, engagement in this research study is conceptually defined as the mobilization of cognitive, affective, and motivational strategies for specific tasks. It is operationally defined as the amount of participants' on-task time (given a viable alternate activity) and the number of starts and completions of the puzzle. Details are described in Section 3.2.1.

2.7 Children's Collaboration

Another variable of interest related to enjoyment and engagement is collaboration. Children communicate and learn through social interaction and imitating one another, and they acquire new skills and learn to collaborate with others. Face-to-face collaboration with classmates or friends is an important part of children's daily life. Numerous researchers have also noted the social and achievement benefits of having children work together in small groups (e.g., Hymel et al., 1993; Inkpen et al., 1995). Collaborative activity is also considered to be a key factor in children's play since it is not only a normal part of play, but also research has repeatedly demonstrated the benefits of children working together (e.g., Inkpen et al., 1999; Stanton et al., 2001).

There has been extensive research on early learning that is organized as an interaction among peers. This interest follows from claims made within influential theories of cognitive development. Learning collaboratively can provide an environment to enliven and enrich the learning process (Piaget, 1928). Piaget (1928), in his early work, did sketch out a role for the significance of interaction between peers. He suggested that young children benefit from peer-based learning because a natural egocentrism necessarily gets challenged – through "the shock of our thought coming into contact with that of others" (Piaget, 1928). Later, he pointed out that collaborative learning has a major role in constructive cognitive development (Piaget, 1932). His influential socio-cultural theory was inspired by other popular learning theories (e.g. Vygotsky, 1978; Thomas et al., 1990) in emphasising the importance of collaboration. The promise of collaborative learning is to allow students to learn in relatively realistic, cognitively motivating and socially enriched learning contexts with the help of collaborating partners. Vygotsky (1978) suggested that cognitive development depends upon either "adult guidance" or "collaboration with a more capable peer". These theories have inspired later empirical research that evaluates peer-based learning in terms of its outcomes. Crook (1997) concluded that three lines of interpretation that mediate an advantage for working collaboratively have emerged. The first dwells upon the fact that collaborators will usually articulate their thoughts publicly (Hoyles, 1985; Schunk, 1986). The second line of interpretation stresses the productive value of conflict that can arise as partners negotiate a consensus (Doise et al., 1984). The third interpretation stresses the possibility of co-constructions within collaborative problem solving (Forman, 1989).

In the field of educational research, there has been some interest in how collaborative activity may facilitate certain kinds of cognitive operations for children. For example, some experimental research has considered whether discussion helps children to generalize what they have learned (i.e., Edwards et al., 1987; Hatano et al., 1992). Hatano (1992) indicated that through sharing ideas, children can achieve a more generalizable understanding if they are actively helped and encouraged to do so. Mercer (1996) discussed the quality of talk in children's collaborative activity.

Research in children's interaction design under a collaboration circumstance has also been well studied. Studies have shown motivational and learning benefits for co-located interaction in computer environments. Inkpen et al. (1999) found that children exhibit a significantly higher level of engagement and activity when working alongside each other. Scott et al. (2003) conducted a study with pairs of children under three collaborative conditions. He indicated that children rather enjoy technology that supports concurrent activities, and working together in small groups increases children's enjoyment, engagement and motivation. Sluis et al. (2004) claimed that a collaborative environment is more likely to elicit increased intrinsic motivation. Based on a review of these claims and the assumption that a collaborative, co-located condition is ecologically valid and would enhance children's enjoyment and engagement for all interface styles, a paired collaboration situation was chosen for this study design.

The rich information available in co-located collaborative environments has spurred researchers to find novel ways of supporting multiple people working together. Different collaboration styles were used by users under different collaborative conditions. There are two collaboration styles I am interested in exploring further in this study. They are the sequential turn-taking strategy and independent parallel play.

2.7.1 Independent Parallel Play

Parallel play is often viewed as characteristic of a "stage" through which children pass as they develop from solitary players to social players. This view is based on Parten's (1932) classic study of children's social participation. Developmental psychology defines parallel play as children playing side by side without interaction (Parten, 1932). In education, parallel play also describes activities where children are divided into pairs or small groups and work on the same activity simultaneously. This gives all children equal opportunity for active involvement (Scarlett, 2004). The conceptual definition outlined above is adopted for this study. Some researchers (e.g., Bakeman et al., 1980; Rubin 1976) have found that children engaged in parallel play may prefer the company of other children, while those successfully considering other points of view engaged in associative or cooperative play (Rubin et al., 1976). Brophy (1974) found that solitary play was often educative, goal directed, independent, and task-oriented in nature. Rubin et al. (1976) suggested that young children chose different levels of social participation based on their cognitive skills; however, young children's sociability may be related to their cognitive style.

In this study, independent parallel play behavior is identified when pairs of children are observed solving at the same puzzle task without verbal, visual or gestural communication.

2.7.2 Sequential Turn-Taking

The turn-taking approach is firstly defined in the field of social science. One definition of this term is proposed by Goffman (1981), i.e., that turn-taking is a form of communication behaviour in which people alternate messages from one person to the other.

In the HCI domain, turn-taking is usually defined as an approach where pairs or small groups of users do a similar task one person after another (e.g., Scott et al., 2003).

Traditional desktop environment currently offer only limited support for concurrent collaboration with peer users. One solution to this limitation is to have pairs of child users share the computer with each other. Scott et al (2003) indicated that the one-mouse-one-cursor accessing paradigm often forces children to interact sequentially. He also indicated that users interact concurrently when the collaborative medium supports it, a capability not offered by typical desktop computer; interestingly, children resisted surrendering the mouse to their partners in the one-mouse setting, even during idle periods (Scott et al., 2003). Inkpen et al. (1997) conducted a study to explore how children's mouse sharing patterns affect their learning in a collaborative environment. She indicated that boys and girls interact quite differently using the various turn-taking protocols.

In this study, sequential turn taking behavior is identified when we observe that only one child in a pair is interacting directly with the puzzle at one time, and as the session progresses the children take turns working one at a time as they solve the same puzzle task.

2.8 Gender

Much empirical research–as well as interaction design research–about children finds that boys and girls like different things, act in different ways, and have differential success at various tasks. Numerous researchers have observed that gender differences often exist with respect to interactions with computers, especially for children (e.g., Cassell et al., 1998; Inkpen et al., 1994). Some previous studies indicated that there is an apparent designer bias towards male users and a corresponding interest imbalance tending towards males being more interested (Cassell et al., 1998; Hughes, 1991). Girls and boys think about computers differently (Hall et al., 1991; Wilder et al., 1985), have different motivations for using computers (Inkpen et al., 1994), and have different preferences (Inkpen et al., 1994; Lawry et al., 1995).

In more general games research, psychologists have discussed that for young children, the organized games of girls are simpler in their rule structure than are the games of boys, and require a lesser amount of physical skill (Cassell et al., 1998; Vail, 1997). Inkpen et al. (1995) observed gender differences in children's puzzle solving tasks under different collaborative interactions. They also noted that boys took and girls relinquished control of the mouse under a collaboration condition (Inkpen et al., 1995).

As a result, it is important that research on children's interactions with computers should be sensitive to gender differences. Thus, gender was one of the factors explored in this study to better understand its impact on children's enjoyment, engagement and collaboration with different interface styles.

2.9 Research Motivation

There has been a growing body of research into approaches for linking the physical and digital worlds. Notable areas include ubiquitous computing, augmented reality, and computer-augmented environments, which have spurred continuing research efforts throughout the 1990s. Simultaneously, a new stream of interface research has begun to explore the relationship between physical representation and digital information, highlighting kinds of interaction that are not readily described by existing frameworks, and exploring the potential benefits to enhancing user's experience with the system.

With the developments in educational technology, the possibilities of incorporating tangibles into learning experience are being explored and introduced. The attempts to adopt this technology in early childhood education, however, have not been devoid of controversy. Despite a growing literature in tangible interaction with children, much of the research in this area has focused on describing the design and technological implementation of tangible prototypes and presenting observations from small user studies (e.g. Africano et al., 2004; Raffel et al., 2006; Verhaegh et al., 2007). This research agenda is grounded in implicit assumptions that tangible style interfaces, which rely on direct physical manipulation and support face-to-face collaboration, are more "natural" and thus more enjoyable and engaging for children than desktop environments. TUIs are assumed to enhance the learning and development process and make children more comfortable with using this technology (Rauterberg et al., 1998; Antle, 2007; Marshall, 2007). However, few empirical studies have addressed the validity of these claims.

Compared to graphical style desktop systems there has been little research which empirically and systematically explores the advantages of tangible systems. It is unknown how the properties of tangible interaction will contribute to enjoyment and engagement in tangible games for school age children. Hence, the claims of the benefits of tangible interaction remain speculative. Understanding these issues will contribute to grounding this technology agenda in empirical studies; inform the development of a stronger framework for the theory and practice of play-based learning with tangibles; and lead to the development of principles to guide the design of new forms of tangibles.

In the design of this research study, a spatial game was chosen because a tangible user interface is a logical choice for spatial activities, especially for children users. Jigsaw puzzles were chosen as they represent a familiar playful learning activity that is undertaken socially, requires cognitive effort, utilizes physical manipulation and is spatial in nature.

In this study, I conducted a controlled explorative experiment to study how interface style affects school-age children's enjoyment and engagement when solving a jigsaw puzzle under a collaborative condition.

2.10 Research Questions

The main research question I explore in this study is:

Does interface style (e.g. tangible user interface, graphical user interface, and traditional user interface) have an effect on school-aged children's enjoyment and engagement when solving a jigsaw puzzle under a collaborative condition?

Tangible user interfaces are hypothesized to be able to positively support children's enjoyment and engagement when solving a jigsaw puzzle rather than other type of interfaces. Individual difference (e.g. gender) may lead to different user preferences relative to the interface styles, which may or may not affect user's enjoyment and engagement.

Detailed research questions are explored and listed as following:

1. *Does interface style affect children's enjoyment on a spatial puzzle task under a collaborative condition?*
2. *Does interface style affect children's engagement on a spatial puzzle task under a collaborative condition?*
3. *Does interface style affect children's collaboration style on a spatial puzzle task under a collaborative condition?*
4. *Does gender composition of pairs affect children's enjoyment differently depending on interface style?*

5. *Does gender composition of pairs affect children's engagement differently depending on interface style?*

6. *What other (demographic) factors affects children's enjoyment or engagement differently depending on interface style?*

The design of this research study will consider all these questions. This thesis presents a comparative study exploring how interface style related interaction factors impact enjoyment and engagement in jigsaw puzzle games under a collaborative condition. My study will contribute to the knowledge base by exploring the influence of tangible interface and interaction features for "spatial" games on engagement and enjoyment of school age children.

In the next Chapter (Chapter Three), I will describe the experimental framework used in this research, starting with the five key hypotheses that address the research questions I posed above.

CHAPTER 3: EXPERIMENTAL DESIGN – FRAMEWORK

3.1 Overview

As discussed in Chapter One and Chapter Two, much of the research with tangible interfaces has focused on development and descriptive analysis of new tangible systems (e.g., Africano, 2004; Raffle et al., 2006). However, little work has explicitly and systematically explored the advantages of tangible user interfaces compared to the other user interface styles (Antle, 2007; Marshall, 2007). This is the only known study to investigate the effects of interface styles on school age children's enjoyment and engagement in a playful learning task.

To investigate how interface style impacts children's enjoyment and engagement, I designed an experimental comparison of school-aged children's enjoyment and engagement on three interfaces for solving the jigsaw puzzles. As mentioned previously, jigsaw puzzles were chosen as they represent a familiar playful activity, which is undertaken socially, requires cognitive effort and physical manipulation and is spatial in nature.

In this research study, subjects were asked to perform a puzzle-solving task in one of three different puzzle implementations in order to explore how interface style impacts their enjoyment and engagement levels. In this chapter, I outline the framework of the experiment design. I describe the research hypotheses (section 3.2), the puzzles (section 3.3), puzzle features and implementations (section 3.4), task set (section 3.5) and measurement tools (section 3.6) for evaluating users' enjoyment, engagement and their relevant collaborative interactions. Data were collected in different forms by both quantitative and qualitative research methods.

3.2 Research Hypotheses

Three different styles of interfaces were examined to determine how they impacted children's enjoyment and engagement. In this study, five main hypotheses were examined for determining the relationship between interface style and the multiple measurements of children's enjoyment, engagement and collaboration.

Enjoyment and engagement are two pivotal dependent variables that were evaluated in this thesis work. The conceptual definitions of enjoyment and engagement set the scope and meaning of the terms. Each is a complex construct derived from physical, social and cognitive theories. Numerous researchers have investigated how children's collaborative interactions can contribute to their enjoyment and engagement on tasks (e.g., Inkpen et al., 1999). Hence, collaboration was considered to be an important variable and observed in this study. The test result was used to contextualize the finding of user's enjoyment and engagement.

The first hypothesis examines how subjects' subjective rating of enjoyment on puzzle tasks is affected by interface style. The second hypothesis examines how subjects' time related evaluation of engagement is related to interface style. The third hypothesis examines how subjects' collaborative behaviour is related to interface styles. Since gender effect is normally discussed in the discipline of children's research, this latent factor will be considered for further examination in hypothesis four and five. The fourth and fifth hypotheses examine how gender composition of pairs affects children's enjoyment and engagement on different interfaces. Other demographic factors (e.g., computer experience) were also examined.

3.2.1 Hypothesis One

Hypothesis One: Interface style will have a significant effect on children's enjoyment on a spatial puzzle task.

I hypothesize that subjects' self-reported rating of enjoyment will decrease from TUI to PUI, and then to GUI. The assumption is based on the claims that tangible style interfaces, which rely on direct physical manipulation and support face-to-face collaboration, are more "natural" and thus more enjoyable for school-aged children than traditional desktop environment. Furthermore, the novelty of the tangible user interfaces will demonstrate its advantage and thus make its more enjoyable for children than the traditional physical user interfaces.

Because *Perceived Competence* is predicted to be positively correlated to enjoyment, subjects will feel most competent while solving the TUI puzzle, but feel least competent while solving the GUI puzzles. Since *Perceived Choice* is predicted to be positively correlated to enjoyment, subjects' agreement of their *Perceived Choice* will decrease from TUI to PUI, and then to GUI condition. *Pressure and Tension* is theorized to be a negative predictor. The score result will be reversed to that of the other subscales. Meanwhile, *pressure and tension* is expected to be negatively correlated to enjoyment. Hence, subjects will rate an increase level of pressure from TUI to PUI, and then to GUI condition.

This hypothesis can be formulated into inequalities as follows:

H_1: Enjoyment: $\mu_{TUI} > \mu_{PUI} > \mu_{GUI}$

Perceived Competence: $\mu_{TUI} > \mu_{PUI} > \mu_{GUI}$

Perceived Choice: $\mu_{TUI} > \mu_{PUI} > \mu_{GUI}$

Pressure/Tension: $\mu_{TUI} < \mu_{PUI} < \mu_{GUI}$

3.2.2 Hypothesis Two

Hypothesis Two: Interface style will have a significant effect on children's objective task time indications of engagement on a spatial puzzle task.

I hypothesize that subjects will spend longest on their first completion of the GUI puzzle, but will spend shortest on TUI condition, subjects' subsequent play time will be longest in TUI condition, but shortest in GUI condition, and subject's amount of off-task time will be highest in GUI condition, but lowest in TUI condition. This is based on the assumption that tangible style interfaces, which rely on direct physical manipulation and support face-to-face collaboration, are more "natural" and thus more engaging for school-aged children than the desktop environment. Direct interaction (i.e, physical or tangible user interfaces) with puzzle pieces is predicted to be easier for children than indirect interaction (i.e., graphical user interface) to complete the puzzle task. Thus, children will take shorter time on their first completions on PUI and TUI conditions than on GUI condition. The novelty of tangible user interface will make the system more engaging and much easier for children to complete the task than the traditional physical user interfaces.

This hypothesis can be formulated to inequalities as follows:

H_1: Time of 1st completion: $t_{TUI} < t_{PUI} < t_{GUI}$

Subsequent play time: $t_{TUI} > t_{PUI} > t_{GUI}$

Off-task time: $t_{TUI} < t_{PUI} < t_{GUI}$

3.2.3 Hypothesis Three

Hypothesis Three: Interface style determines children's collaboration style on a spatial puzzle task.

Children's collaboration styles on tangible user interfaces are predicted to be similar to that on physical (traditional) user interfaces, but different to that on graphic user interfaces. I assume that pairs of children will predominantly use the independent parallel play strategy on both physical (traditional) user interfaces and tangible user interfaces, but use the sequential turn-taking strategy on graphical user interfaces. This assumption is based on the conclusion given by recent research studies, which claimed that tangible or physical user interfaces can provide more space to contribute to users' face-to-face collaboration, especially for children (e.g. Africano et al., 2003; Fails et al, 2005; Antle 2007; and Marshall, 2007).

3.2.4 Hypothesis Four

Hypothesis Four: Gender composition of pairs will have a significant effect on children's enjoyment depending on interface styles.

In research on children's play, there is a lot of discussion about gender difference. Therefore, gender composition of pairs will be examined to see whether it has an effect on subjects' enjoyment depending on different interface styles.

Based on the large literature reviews of children's gender effect and their different attitudes to new technologies (e.g., Cassell et al., 1998), boy-boy pairs are predicted to enjoy computational and novel technology-embedded games more than the girl-girl pairs. I hypothesize that enjoyment level on the TUI condition will be higher than the other two kinds of conditions evaluated by all gender composition of pairs. This is in line with the assumption that I used for hypothesis one, in which I indicated that tangible style interfaces, which rely on direct physical manipulation and because of the novelty factor, are more enjoyable for school-aged children than other two kinds of interface styles. Boy-boy pairs should enjoy the TUI puzzle the most, but the PUI puzzle the least. Mixed gender pairs are predicted to offset the differentiations in between. However, girl-girl pairs are hypothesized to prefer traditional cardboard puzzle rather than the GUI puzzle.

3.2.5 Hypothesis Five

Hypothesis Five: Gender composition of pairs will have a significant effect on children's engagement depending on different interface styles.

The evaluation of engagement was continued by examining whether gender composition of pairs has an effect on the result. Inkpen et al.'s (1995) study indicated that when asked to play a puzzle game on computer under an integrated play condition, girls solved fewer puzzles than boys did. Thus, girl-girl pairs are predicted to use more time on their first completion on any kind of interfaces than boy-boy pairs. Theory and research have suggested that the experience of enjoyment is a critical factor in determining one's motivation for and continued participation in exercise settings (i.e., engagement)(Kremer et al., 1997; Wankel, 1993). Therefore, boy-boy pairs are predicted to stay much longer for their subsequent play in the TUI and GUI conditions than the girl-girl pairs. Mixed gender pairs are predicted to offset the differentiations in between. To be in line with the previous assumption of children's engagement on different interface styles I indicated in hypothesis two, the subjects' overall on-task time under the TUI condition is predicted to be much longer than that of the GUI condition.

3.2.6 Demographic Variables

User's demographic information was collected through a pre-questionnaire (Appendix B). The effects of individual difference were controlled as much as possible. Where they couldn't be controlled, they were examined to determine if they affected subjects' enjoyment and engagement level significantly in further analyses. These include subject's age, gender, native language, computer experience, competence of mouse control, puzzle experience, preference of puzzle, and preference of themes.

3.3 The Puzzles: PUI, GUI and TUI

All puzzles used one of two different content themes, each with the same modern style of cartoon illustration. One theme was a whimsical illustration of an imaginary castle with bats, ghosts, witches, knights and a princess. The other theme was an illustration of the legendary pirate Barbarossa and his ship, the Black Pearl (see Figure 3.6). Both themes are inclusive of gender and are currently popular in children's media as can be seen in the success of the Harry Potter and Pirates of the Caribbean books and movies.

Figure 3.1 Two puzzle themes

3.3.1 Traditional Style (PUI) Puzzle

The two cardboard jigsaw puzzles chosen for the experiment were designed and manufactured by DJECO, a European game publisher. Each puzzle consisted of 54 pieces (6 x 9). The dimensions of the completed puzzle were 42 x 45 centimetres. Both puzzles were recommended for children older than 5 years. In a pilot test, we determined that this size puzzle could be completed by two six year olds in fifteen minutes. Each puzzle came with a poster of the image which we used as the underlay for the puzzle.

3.3.2 Graphical Style (GUI) Puzzle

The two GUI puzzles were created using commercially available jigsaw puzzle creation software, "Jigs@w Puzzle 2", developed by TIBO software. Each puzzle was run on a laptop with Intel Core 2 Duo Processor, a 15.4" (39.1cm) wide-screen WXGA display, and equipped with a Microsoft wired optical mouse. The game interface occupies the full screen of the computer (see Figure 3.7). The puzzle pieces were manipulated by using drag-and-drop technology. Each piece could also be rotated right-clicking the mouse. Users could select to either show or hide a real size reference picture in the background. When pieces were correctly connected, they were connected permanently. Visual and audio feedback was provided by the software for correct matches. I found through the pilot study that the size of puzzle piece displayed on the laptop screen was smaller than the physical piece, and it affected participants' average completion time (discussed in Section 4.4). I adjusted the total number of GUI puzzle pieces to be 42 pieces (6 x 7) to address this problem and ensure that the three implementations were of comparable difficulty.

Figure 3.2 Screen shot of the GUI puzzle

3.3.3 Tangible Style (TUI) Puzzle

The two TUI puzzles were implemented on two identical, extensible tabletop prototypes designed specifically for this study (Figure 3.8). The puzzle pieces were two new versions of the traditional version. Input actions on puzzle pieces were captured using an infrared web camera embedded under the table. The ReacTIVision engine was used for fiducial marker recognition (Jordà et al., 2007). However, instead of marking each individual puzzle piece with a unique fiducial pattern, the markers were distributed along the edges of intersecting pieces. No one puzzle piece had an entire pattern (Figure 3.9). Instead, the system recognized user triggered events, which are when a correct connection between two or more pieces was made with the physical pieces. When pieces are assembled & placed correctly, a location-based, puzzle shaped image was blacked out on the reference image, which was projected on the surface of the tabletop system. Meanwhile, the system also played a laser sound effect to let user know their proper connection of the puzzle pieces. Figure 3.10 illustrates a sample of two correct puzzle pieces connected. The connection made a complete fiducial marker at their connected edges. In response to these input events, a logic program, implementing with Processing, was used to control visual and audio feedback similar to the GUI feedback. The puzzle images are projected and reflected onto the surface of the table via a piece of mirror underneath the table. The system provides a projection area of 42x45 cm (at a resolution of 1024x768 pixels). The physical dimension of the tabletop system is 92x76x61 cm (LxWxH). The size of the table is specially designed for children, on which they can take part in a comfortable and convenient interaction. The final prototype was a tangible interface to the physical jigsaw puzzle that embodied the properties and functions of both the PUI and GUI.

Figure 3.3 The prototype of the TUI puzzle

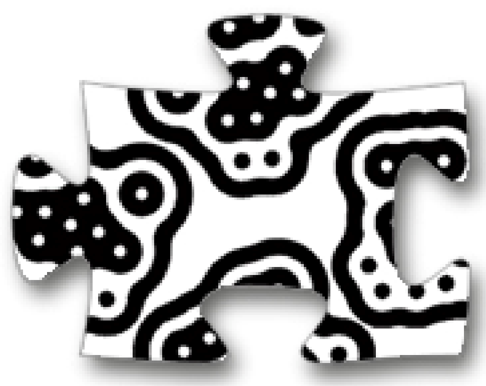

Figure 3.4 A sample of one puzzle piece

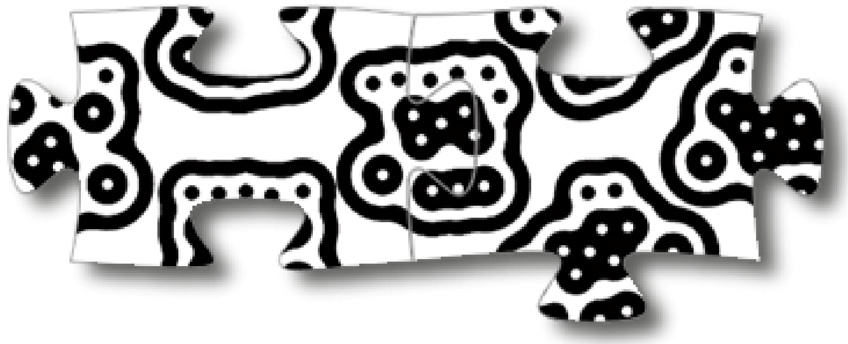

Figure 3.5 A sample of two connected puzzle pieces

3.4 Puzzle Features and Functions

To facilitate a valid comparison, I used the same two puzzles implemented in each interface style. In the initial design, I held many of the physical characteristics of the puzzle constant across all three implementations, including: image style, visual content, size of puzzle piece, the number of pieces, and availability of underlying image.

The key differences among the three implementations are related to modality of feedback and styles of social and physical interaction. This study design enables the investigation to focus on the features of TUIs which are often cited as enjoyable and engaging: direct physical manipulation (PUI, TUI), integrated image reference (PUI, TUI), face to face social interaction (PUI, TUI), and integrated feedback (GUI, TUI) with the PUI acting as a control. Table 3.1 indicate the differences and similarities of the three puzzle implementations on these four themes.

	PUI	GUI	TUI
Direct physical manipulation	+	–	+
Integration of reference	+	–	+
Face-to-face collaboration	+	–	+
Digital auditory & visual feedback	–	+	+

Table 3.1 Differences and similarities of three puzzle implementations

The traditional puzzle lacks digital auditory or visual feedback. However, children can directly manipulate the puzzle piece and received face-to-face social interaction. A poster of the puzzle underlies the traditional puzzle to serve as integrated paper-based visual reference. The GUI puzzle involves indirect manipulation by a single user (via the mouse or touchpad), the degrees of freedom of movement of puzzle pieces are limited to two dimensions (rotation in 2D is possible). The puzzle size is limited by display size. The system has a separation of the input and output image. But, it provides the digital auditory and visual feedback to its users. The TUI puzzle shares the style of direct physical interaction in three dimensions and the possibility of face-to-face social collaboration with the traditional puzzle. It has the integrated reference image, and was implemented to include the same modalities of feedback (auditory and visual) and available operations as the GUI puzzle (e.g., turning the underlying image on/off; reseting the puzzle).

3.5 The Task

The experimental design was a three-by-two, fully balanced with interface style (PUI, GUI or TUI) and puzzle themes (pirate boat, wizard & witch castle), for a total of six treatments. The study task is considered to be equally represented in the three different interface styles. In order to eliminate the order effect, each set of subjects was asked to play only one puzzle on only one of the three interface styles.

3.6 Measures

This study design facilitated the collection of several forms of quantitative and qualitative data. Questionnaire data was collected for analyzing subjects' enjoyment level. Time-related data was collected for evaluating subjects' engagement level. Behavioural-based observational data of subjects' interaction was used to explore their collaborative interaction.

3.6.1 Enjoyment

Since we are interested in children's social play, a conception of enjoyment based on intrinsic motivation is relevant. The Intrinsic Motivation Inventory (IMI) is a validated multidimensional measurement instrument based on SDT (Ryan, 2006). It was designed to measure participants' subjective experiences related to *Interest and Enjoyment* in activities conducted in laboratory experiments by measuring intrinsic motivation. It takes the form of a questionnaire utilizing a Likert scale. In addition to measures of *Interest and Enjoyment*, it has five related subscales. Three of the subscales were pertinent for this study, and they are *Perceived Competence*, *Perceived Choice* and *Pressure and Tension*. The questionnaire contains seven questions related to the *Interest and Enjoyment* subscale, five questions related to the *Perceived Competence* subscales, five questions related to the *Perceived Choice* subscale, and five questions related to the *Pressure and Tension* subscale. The validities of these four subscales have been established across a variety of tasks, conditions, and settings (Ryan, 2006).

As discussed in previous chapter, enjoyment in this study is operationally measured through subjective ratings of these four subscales of IMI questionnaires. *Perceived Competence* is predicted to be positively correlated to enjoyment. *Perceived Choice* and *Pressure and Tension* are included to provide a measure of the impact of the artificial nature of a lab study. *Perceived Choice* is predicted to be positively correlated to enjoyment. Perceived *Pressure and Tension* is expected to be negatively correlated to enjoyment. In order to avoid bias, the questionnaires were modified to be suitable for being used by all subjects no matter what kind of interfaces they were tested on.

A post-questionnaire, based on a modified version of the four subscales of the Intrinsic Motivation Inventory (IMI) (Ryan, 2006), was given to the participants after each session. Instead of using the seven-point Likert scale that the IMI questionnaires originally have, we used a five-point rating scale based on the Smileyometer, which has been validated for collecting children's subjective ratings (Read et al., 2002). The scale uses a pictorial representation of five different smiley faces as shown in figure 3.9. Subjects were asked to circle one of the faces for demonstrating the truth level of each statement in the questionnaire.

The five values are presented to be:

- Not at all true
- Not very true
- Somewhat true
- True
- Very true

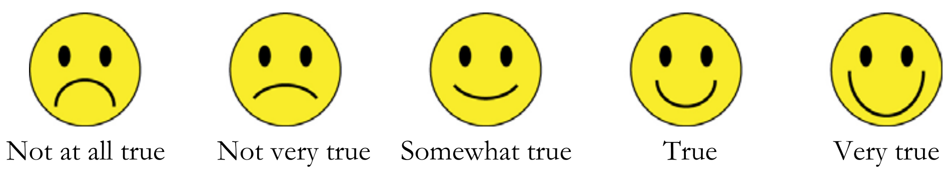

Figure 3.6 Smileyometer Scale

A coding scheme based on the original point-scale design of IMI was used for coding the illustrated facial representations. For data analysis, "Not at all true" was coded as one and "very true" was coded as five.

In addition, I concluded the sessions with two additional open-ended questions related to participants' preferences during their experience with the puzzle. The results were used to contextualize the quantitative findings of enjoyment on different interface styles.

3.6.2 Engagement

Instead of using subjective observation of facial expression or behaviour, engagement in this study is operationally defined as the amount of participants' on-task activity time (given a viable alternate activity) and the number of starts and completions of the puzzle. Figure 3.7 illustrates three components of subjects' activity time and the relationship between them.

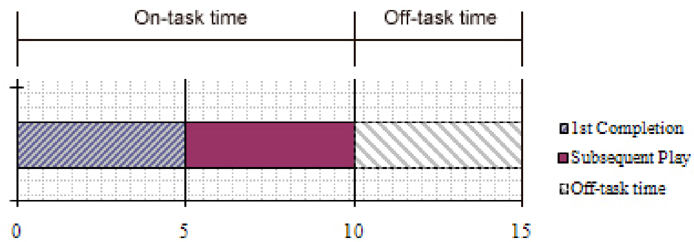

Figure 3.7 Time related measures of engagement (on-task time vs. off-task time)

During each session, each set of subjects was given the same amount of time (15 minutes) to play with the puzzle. During the entire allocated time, we recorded the time they used for their first completion, their subsequent play time (if they stayed for repeat plays) and the off-task time (if they left for other activities). User's on-task time was operationally defined as the time user spent on interacting with the puzzle implementations. Conversely, off-task time was counted as user's time expended on irrelevant tasks (e.g. subjects quit the puzzle and went to the book station instead). I also developed a equation indicated the detailed relations among allocated time, on-task time and off-task time by calculating the time of users' first completion, time of their subsequent play time and the off-task time.

$$T_{\text{allocated time}} = t_{\text{1st completion}} + t_{\text{subsequent play}} + t_{\text{off-task time}}$$

During each session, observational notes were recorded related to task time. For example, the facilitators recorded if one member of a pair quit before the other. We also counted the number of times pairs began and the number of times they completed the puzzle in the allotted 15 minutes. All sessions were video taped for validation of time.

3.6.3 Collaboration

Another variable of interest related to enjoyment and engagement is collaboration. Collaboration in this study is primarily analyzed from qualitative findings. The qualitative analysis was based on informal observation and digitized audio recordings of the two post-play open questions. Quantitative analysis based on subject's responses to the two collaboration-preference-related questions enclosed in post-questionnaires is also used.

During each session, the facilitators took observational notes. The observations focused on classifying subjects' dominant collaboration style as either independent parallel play or sequential turn-taking play. As discussed previously, independent parallel play is operationally defined as pairs of children work independently without verbal, visual or gestural communication. Sequential turn taking is operationally defined as pairs of children taking turns sequentially to work on solving the same puzzle task. Other kinds of children's behavioural-based indications of collaboration were also observed and examined in the user study, such as whether or not pairs of children work cooperatively, and whether or not they work competitively.

3.6.4 Demographical Variables

Each session starts with a pre-questionnaire (Appendix B). The pre-questionnaire was designed to collect participants' demographic information, computer experience level, and preference on jigsaw puzzles and image themes. These factors were further examined to determine if they affected users' enjoyment, engagement or collaboration significantly.

In Chapter Four, I outline the experimental methodology used to examine the relationships with interface style and subjects' enjoyment, engagement and collaboration in the design of this study.

CHAPTER 4: EXPERIMENTAL DESIGN – METHODOLOGY

4.1 Overview

In order to test the five hypotheses I outlined in Chapter Three, a controlled comparative empirical experiment was devised. The experiment examined subjects' interaction with different user interface styles for solving a puzzle. Pairs of subjects were asked to play one puzzle on one of the three interface styles in the main experiment. A one-factor between-subject design was conducted, with the main factor being interface style (i.e., physical, graphical and tangible user interfaces). The study continued by involving a secondary factor, which is gender pairing. A two-factor between-subject study was conducted afterwards. The study also furthered its exploration by analyzing the affect of other covariates, such as computer experience or preference on puzzle/puzzle themes.

One pilot study and one subsequent main experiment were conducted. Information gathered from the pilot study was used to fine-tune the design of the subsequent main experiment. The main experiment began with a pre-questionnaire (as discussed in Chapter Three). Subjects were arbitrarily grouped into pairs when participated in the activities. The pairs were asked to complete one kind of puzzle on one kind of interface within an allocated amount of time. Then the pairs were asked to complete a post-questionnaire which was verbally administered. The session concluded with two open-ended questions related with subjects' preference on the puzzle they just played.

In this chapter, I describe the participants (section 4.2), study setting (section 4.3), assumptions (section 4.4), the pilot study (section 4.5), the design of main experiment (section 4.6), and the statistical analysis tools (section 4.7) and qualitative analysis methods (section 4.8) I used for this research study.

4.2 Participants (Subjects)

There were 140 children in total who participated in this study. Eight children participated in the pilot study and 132 children took part in the main experiment. All the participants were recruited without any discrimination other than satisfying the age constraint (6-10 years) and being fluent in English. One of the reasons that we constrained our participants in this specific age group is to ensure the formation of a relatively homogenous group, thus minimizing some of the effects of age difference and validating the generalizations. For example, it was assumed that subjects would have a similar capability of completing the puzzle tasks and a similar understanding of all the questions we presented in the study.

In the pilot study, all subjects were recruited from Symphony Safari Summer Camp at SFU Surrey. Subjects were selected based on the age criteria. The eight children were randomly grouped into four pairs. Each pair was asked to complete two puzzles on two different user interfaces sequentially.

The main experiment was conducted at Science World, Vancouver, B.C. All the 132 participants were recruited from visitors. They were recruited from the regular visitor population using posters in various locations at the centre (outside the lab) or from ongoing centre summer camp participants. Participants were arbitrarily grouped into pairs depending on recruiting sequence. Pairs were assigned to one of the three different interface styles without any preference. Children were recruited to do a "Puzzle Study" and did not know about the different interface styles before volunteering.

4.3 Setting

The pilot study took place in the EC3 lab (Room 3930) at Simon Fraser University Surrey. In order to facilitate the data collection, two systems and facilitators were provided simultaneously. Two separate spaces were setup with each having prepared a tabletop system and a laptop computer. Each facilitator was asked to monitor one pair of participants in a designated space. A shared area with pillows and a collection of popular children's books was setup in the middle of the room. The purpose of placing this area is to provide an alternative for subjects in case they need to fulfil their entire allocated duration with an activity other than solving puzzle tasks. Subjects' time expended interacting with this area was recorded for future analysis as well.

The main experiment took place at Science World during a three-week period in the summer of 2007. Science World is an interactive science museum where children and adults explore scientific concepts through a variety of hands-on activities. We set up our study in the Eureka lab, a partially enclosed lab space which was relatively isolated and allowed for environmental control during the study. The lab was setup differently on different days depending on the three different experimental conditions. A child's size table (comparable to the TUI table) was used as a space for children to fill out questionnaires, and to set up the GUI and PUI style puzzles. A rest area with benches, pillows and a collection of popular children's books was also provided at the rear of the room. The figure 4.1 illustrates the floor plan of the experiment setting in the Eureka lab.

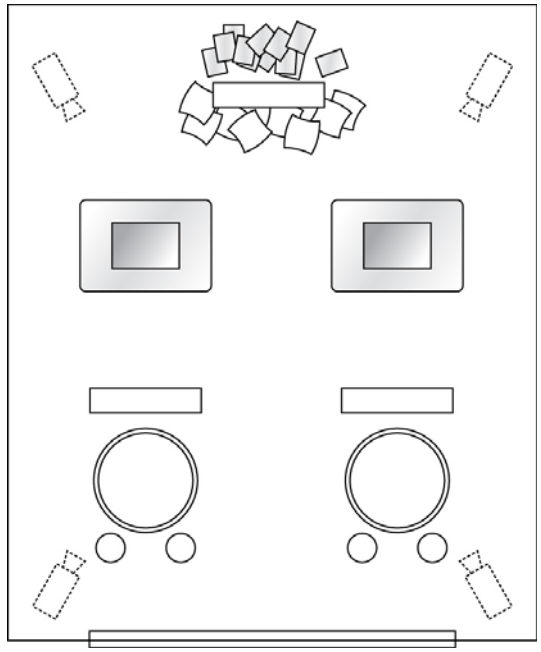

Figure 4.1 The floor plan of the main experiment setting at Science World

4.4 Assumptions

The design of this study assumes that the effect of implementing different interface styles for puzzle tasks can be isolated. This is because a series of physical characteristics of the puzzles, which include image style, visual content, puzzle piece size, the number of pieces, and the availability of underlying image, were kept the same across all interface styles.

The subjects' performance in this study reflects a mix of many factors, including subjects' computing experience, their familiarity with the mouse control, their preference with the puzzle, puzzle theme, images, and their gender, age, and native language skill. Hence, information on all these factors was collected via a pre-questionnaire as part of subjects' profiles in order to determine if these factors had an effect on subjects' performance. Also, a relatively large sample size for this kind of study helped to average out effects of individual differences.

Since the TUI puzzles were developed from the sturdy cardboard puzzles, and both of them shared the exactly same physical characteristics, it is assumed that the difficulty level of completing these two puzzles were similar to the users. This was a reason why the PUI puzzles were not tested in the pilot study. The GUI puzzle used for the pilot study consisted of 56 pieces (7×8), which was the most similar number of pieces the puzzle software could generate. In terms of the task achievability difference of the TUI/PUI and GUI puzzles, it was assumed that the GUI puzzles were more difficult than the TUI/PUI puzzles. This is because of the smaller display size of each puzzle piece, the indirect manipulation of the digital puzzle pieces and its single user mode, and thus subjects might require more time to complete the task. This was seen in the pilot study. The results indicated a big difference in subjects' completion times between TUI and GUI puzzles. I decided to adjust the total number of GUI puzzle pieces to be 42 pieces (a design of 6×7) to minimize the achievability difference across all the three types of puzzles and ensure that the three implementations were comparable.

In terms of the order effect, it was assumed that the sequence of the puzzle play on different interface styles might affect subjects' rating of enjoyment level and their engagement level with the tasks. This was verified in the tests of the pilot study, as the results indicated that most children demonstrated losing their patience and interest in puzzle playing in the second session. Their report of their enjoyment and engagement in this session showed a big difference compared to the first session. Thus, the main experiment was designed to conduct only one session with each pair of participants.

4.5 Pilot Study

The main purpose of conducting the pilot study was to fine-tune the design of the main experiment. Eight children aged from 5 to 8 were involved in the pilot study. Subjects' performance, their feedback on the systems, question wording and average completion time were examined. The technical performance of the systems was also tested. Issues found in the pilot study were tuned in the design of main experiment.

In the pilot study, I test two of the three interfaces, the GUI and TUI, with 8 children. Subjects were grouped into pairs for completing the tasks. Each pair was asked to test on two interfaces sequentially. Five main elements were examined, and they are:

- systems technical performance and its stability
- subjects' average task completion time on each interface
- subjects' understandability of wording (in both questionnaires and script)
- data collection process (including tool functionality)
- overall experimental flow

The entire duration allowed for each pair was one hour. The study consisted of two sessions. Each pair used one interface (GUI or TUI) with either castle-theme or pirate boat-theme puzzle, for a session. They then played on the other interface style with the other theme of the puzzle in the second session. To counterbalance the order effect, two teams were assigned to start with a TUI puzzle, and then they were asked to complete a GUI puzzle in the second session. The other two teams started with a GUI puzzle first followed by a TUI puzzle afterwards. Each set of studies was administrated by one facilitator. At the beginning, the participants started with a pre-questionnaire, which was verbally administered to mitigate for variation in reading skills and ensure adequate comprehension of questions. Pairs of children were then shown one puzzle implementation and asked to solve this jigsaw puzzle together. Each pair was told they would have 15 minutes to play with the puzzle. They were told that they could stop playing the puzzle at any time and instead move to the book section. When the 15 minutes was done, the children were asked to complete a post-questionnaire which was also verbally administered. They then were given a five minute break and were told to be ready for their second play session. After the break, they were introduced to the other puzzle implementation and were told to solve the second puzzle together within the same amount of duration. They were also told that they could decide to stop playing at any time and move to the book section instead. After the second play session was done, both of them were asked to complete a post-questionnaire that was verbally administered again. The post-questionnaire was exactly same as the one they used for their first session.

One of the primary interests was the time requirement of completing each task. It was expected that each task would require a similar amount of time, and could be completed by pairs of subjects within the allocated duration (15 minutes). Means of users' task-related times on different interfaces based on eight subjects are listed in Table 4.1.

Interface Style (Sequence)		1st Comp.	Subsequent Playing	On-task Time	Off-task Time	On-task (%)
TUI	TUI->GUI	13.73	1.27	15	0	100%
	GUI->TUI	14.25	0.75	15	0	100%
	AVG	13.99	1.01	15	0	100%
GUI	TUI->GUI	N/A	N/A	8.75	6.25	85%
	GUI->TUI	N/A	N/A	12.75	2.25	58.33%
	AVG	N/A	N/A	10.75	4.25	71.67%

Table 4.1 Means of subjects' task-related times

None of the eight users could complete the GUI puzzles within the total given time. Two teams, which were assigned to start with TUI task and do GUI task afterwards, quickly lost their interest in solving the second puzzles and quit without completion. Furthermore, verbal debriefing comments indicated that all users reported feelings of frustration and confusion when solving the GUI puzzle. Some children complained that the size of the puzzle piece was too small and the number of the puzzle pieces was too much. Hence, the GUI tasks were modified, and the number of pieces was decreased to 42 pieces (a design of 6X7) for balancing the equivalent cognitive difficulty to the TUI and PUI puzzles.

Means of subjects' rating scores of the four subscales of the IMI on different interfaces with different task sequences is listed in Table 4.2. The graph (Figure 4.2) illustrates different means of rating scores of each subscale on different interfaces. The result revealed that sequence might have an effect on users' rating of each subscales on the different interfaces. Verbal debriefing comments also pointed out that the design of the one-hour study seemed too long for young users to complete both sessions (tasks) with equal enthusiasm, interest and motivation. Thus, it would affect users' overall rating on enjoyment and their engagement level on their second task. Because of this, the design of the study was modified.

Interface Style (Sequence)		Interest Enjoyment	Perceived Competence	Perceived Choice	Pressure Tension
TUI-GUI	GUI	2.11	2.00	2.80	3.50
	TUI	4.96	4.80	4.20	1.80
GUI-TUI	GUI	3.89	2.40	3.55	2.30
	TUI	4.82	3.95	3.65	2.55

Table 4.2 Means of rating on four IMI subscales vs. interface styles and task sequence

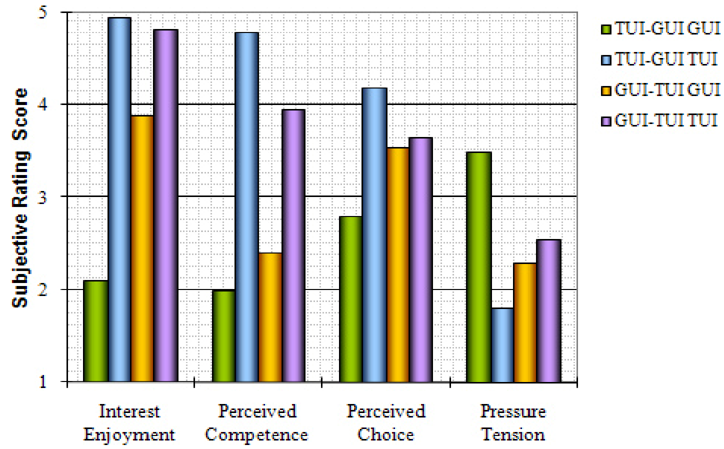

Figure 4.2 Means of four IMI subscales vs. interface style and task sequence

Subjects' performance on the pre- and post-questionnaires revealed their common difficulty in understanding a series of words, which included nervous, tense, relax, anxious, skilled, pressured, enjoyable, and competent. Thus, a script of the interpretations for these words was required for the main experiment (Appendix A: Script).

Modifications based on the results of the pilot study included:

- The total number of pieces of the GUI puzzles was decreased
- The second session of the study was eliminated
- The session script was modified according to the new study design
- Prepared the script of interpretations for a series of words used in the questionnaires.

4.6 The Main Experiment

The main experiment was comprised of sessions with one hundred and thirty two children of both genders, aged predominantly in the 7-9 years range, recruited from the regular visitor population or the on-going summer camp participants at Science World. Participants were arbitrarily grouped into pairs depending on their visiting sequence. No previous puzzle solving experience was required. During all the study sessions, two facilitators were present in the Eureka lab simultaneously. The research assistant who facilitated the pilot study also helped to conduct all the sessions.

The detailed experimental protocol and verbal scripts used for different interface style conditions are listed in Appendix A. The entire duration of each session for a single pair was 30 minutes. Prior to the study, each subject's parent or legal guardian was required to fill out a consent form. At the beginning of each session, the subjects were given a brief orientation. Each participant was given a chance to briefly introduce him- or herself by name. They were then asked to begin with a verbally administered pre-questionnaire.

After the orientation and pre-questionnaire sessions, pairs of children were then shown the puzzle implementation and asked to solve a jigsaw puzzle together. Facilitators would demonstrate the general functions of each implementation to the children before starting. In GUI sessions, facilitators demonstrated how to use mouse to move the puzzle pieces or to spin the pieces, and how to display or hide the reference image on the screen. In TUI sessions, facilitator demonstrated how to make the system display or black out the projected image and explicitly told subjects that the system would respond to the location-based correct connection by playing a laser sound. Participants were also given chances to ask questions before starting the activities. They were then instructed to spend a total of fifteen minutes solving the puzzle together and could stop playing the puzzle at any time and move to the book area instead. When both subjects indicated they were ready, the facilitator would leave children solving the puzzles by themselves and start observation (taking observation notes). Elapsed time, counting and videotaping started at this point too.

After the fifteen minute play session was finished, the children were asked to complete a post-questionnaire, which was verbally administered by the facilitator. The facilitator would verbally explain each question to all the participants. The session ended with a closing audio-taped interview in which the children were asked about their impressions and preferences over the puzzle and what they liked or disliked about it.

The main experiment was designed to test the effect of three interface styles, in terms of children's subjective ratings of *Interest and Enjoyment*, *Perceived Competence*, *Perceived Choice* and *Pressure and Tension*; the effect of interface styles, in terms of users' time and counts related to engagement; the effect of interface styles, in terms of users' collaborative interaction; and the effect of individual differences (e.g. gender). Interaction effects and correlation between variables were also analyzed. The analysis result is given in Chapter Five.

4.7 Statistical Analysis

For data analysis purpose, all answers to the questionnaires were coded into numbers. There are seven questions related to the *Interest and Enjoyment* subscale, five questions related to the *Perceived Competence* subscales, five questions related to the *Perceived Choice* subscale, and five questions related to the *Pressure and Tension* subscale. The sequence of questions representing the four different subscales was shuffled in the questionnaire, and some statements of the questions were reversed (Appendix C). The scores were firstly divided into groups according to the four IMI subscales. For those reversed questions, the item score is retrieved by subtracting from six. Then, I calculated subscale scores by averaging across all of the items on that subscale. The subscale scores were then used for future data analysis. Missing scores were dealt with by substitution with the mean value for each subscale from each user. The process yielded 132 sets of complete data.

Since the study design determined that user's responses was dependent to each other, data collection was based on a pair condition. Therefore, the quantitative analysis employed a majority of pair-based measurements on both enjoyment and engagement evaluations. Scores of four IMI subscales from each unit (pair set) were averaged from each partner's subscale scores. This further process yielded 66 sets of complete pair-based data sets.

The majority of data collected from the pre- and post-questionnaires used the one to five Likert-type scale tool. Likert scaled data are fundamentally rank ordered rather than interval scaled (Gardner and Martin, 2007). Standard statistical tools, such as *t*-tests, ANOVA, and regression assume that numerical scales on which data are measured behave in a regular way. However, Likert scale data based on human's subjective responses is often shows an irregular distribution because subjects' interpretation to the answers can vary widely. Thus, Likert-type data need to be analyzed using rank based statistical procedures (Gardner and Martin, 2007).

A formal test of normality was run based on the scores of the four IMI subscales collected from the 132 subjects (66 pairs) in this study, and the results are presented in Table 4.3. The Shapiro-Wilks W test was chosen because it is recommended for small and medium samples up to n = 2000 (SPSS). The Shapiro-Wilk W test results of the four subscales (*Interest and Enjoyment*, *Perceived Competence*, *Perceived Choice*, and *Pressure and Tension*) are significant ($p<0.05$), which suggests that the assumption of normality is not met. Statistical data distribution charts of enjoyment data illustrate that the shape of the curve is not normal. The normality test run based on the time-related engagement data (on-task time) also indicates that the data was not normally distributed ($p<0.05$) (Table 4.4). The shape of the curve is not normal either. Both enjoyment and engagement data are not normally distributed. Hence, nonparametric test tools were suggested for all the data analyses.

	Shapiro-Wilk		
	Statistic	df	Sig.
Interest & Enjoyment	.837	66	.000
Perceived Competence	.865	66	.000
Perceived Choice	.906	66	.000
Pressure & Tension	.951	66	.011

Table 4.3 Test of Normality on report of the IMI subscales

	Shapiro-Wilk		
	Statistic	df	Sig.
On-task time	.871	66	.000

Table 4.4 Test of Normality on report of on-task time data

To analyze date related to hypothesis one, hypothesis two and hypothesis three as outlined in Chapter Three, the Kruskal-Wallis tests were chosen. This is because the study design has three levels of independent variable. Data analysis requires an alternative of a one-way ANOVA. The Kruskal-Wallis test is a nonparametric counterpart to the one-way analysis of variance. With the Kruskal-Wallis test, there are no normality or equal variance assumptions (Elliott and Woodward, 2007). If a significant Kruskal-Wallis test is obtained, multiple Mann-Whitney tests can be used to examine pairwise differences. The correlation between the *Interest and Enjoyment* subscale and the other three subscales outlined in hypothesis one were examined using Spearman's Rho test. Spearman's Rho test is a nonparametric alternative to correlation analysis.

To test hypothesis four and hypothesis five as outlined in Chapter Three, the Friedman test was chosen. Friedman's test is a nonparametric alternative to the two-way ANOVA. It is chosen for evaluating the combined effect of two experimental variables (factors). The main purpose of using Friedman's test is to examine the interaction effect based on interface style and gender composition in pairs. Gender pairing as one main factor is defined under a pair-based condition. Gender pairing is categorized as following:

- Group 1: Boy and boy
- Group 2: Boy and girl
- Group 3: Girl and girl

In order to explore the possible relations between other demographic variables (most of them were nominal data) and enjoyment or engagement data, multilevel covariance analysis was performed, where we have the fixed independent variable that we manipulated (interface style) and one covariate attached. Variables were defined under both individual and pair-based conditions due to demand. The pair-based variable tested in this study is the age group.

Age group is divided based on the following criteria:

- Group 1: both are younger than 7-year-old
- Group 2: one is younger than 7, the other is 7-9 year old
- Group 3: both are 7-9 year old
- Group 4: one is 7-9 year old, the other is older than 9

Individual variables included: age, native language, computer experience, competence of mouse control, puzzle experience, preference of puzzle, and preference of theme.

Computer experience is categorized as follows:

- Group 1: Never used computer
- Group 2: Using computer once a month
- Group 3: Using computer once a week
- Group 4: Using computer every day

Demographic variables with two or three values included: puzzle experience (yes and no) and native language (yes, no and somewhat). Three variables were coded based on a Likert scale question, and they are: competence of mouse control, preference of puzzle, and preference of theme.

4.8 Qualitative Analysis

The study design facilitated the collection of several forms of qualitative data. The qualitative analysis was based on informal observation and digitized audio recordings of the two post-play open questions.

4.8.1 Observations

Collaboration findings were analyzed based on observational notes. The observational notes were taken by two different facilitators. Observational notes are recorded based on predicted themes of children's collaboration styles. As discussed in section 2.7, there are two collaboration styles that I have predicted will be commonly observed on three interface styles. They are independent parallel play and sequential turn-taking play. Other kind of collaborative interactions, if applicable, such as directive play, cooperative play and competitive play, will be classified as one category. Collaboration style in this study is categorized as:

- Independent parallel play
- Sequential turn-taking play
- Other styles (directive dominant play, cooperative play or competitive play)

As discussed previously, four features of tangible user interfaces, which include direct physical manipulation, integration of input and output space, face-to-face collaboration and digital feedback aspect (Antle, 2007), are important to the design of tangible and spatial interaction for children. Through the informal observations, I also looked for evidence of whether these features impact any of the main dependent variables (i.e., enjoyment and engagement) on different interface styles.

4.8.2 Audio Recordings

Results of children's preference based on their answers to the post-play open questions are used to contextualize the findings of children's enjoyment on different user interfaces. The digital audio recordings of the two post-play open questions are coded into two preference topics, which are the like and the dislike questions. Children's answers to these two questions yield several discussion themes. These themes are included in Table 4.5.

Like	Dislike
Help from the image	Without the help of image
The picture	The picture
The theme/character	The theme/character
Sound	Sound (distraction)
Help from partner	The partner took all turns
Challenging	Too hard/too many pieces
Can be completed	Can't finish
Solving puzzles	Time limitation
Moving puzzle pieces	Mouse control problem
Others	Nothing

Table 4.5 Coding themes for subject's response to preference

In Chapter Five – Results, I provide the complete set of statistical results, as well as the specific details of the statistical analyses used in the main part of this study. Both quantitative and qualitative data was analyzed to test the five hypotheses outlined in Chapter Three. In Chapter Six – Discussion, I discuss all the statistical results and qualitative findings through this study in order to examine the implications of these findings and discuss the generalization to the design of tangible user interfaces for school age children.

CHAPTER 5: RESULTS

5.1 Overview

To test all five of the hypotheses addressed in Chapter Three, statistical analysis was performed on data collected in the main experiment. In this chapter, I present the data analysis outcome and the results obtained. Hypothesis one examined the effect of interface style on subjects' rating of enjoyment and the other three relevant subscales (*Perceived Competence*, *Perceived Choice* and *Pressure and Tension*). Data analysis was continued with the correlation tests between pair wise subscales. Hypothesis two examined the effect of interface style on users' objective time and counts of measures of engagement. The relationship between users' collaborative interaction and interface style as outlined in hypothesis three was explored using both qualitative and quantitative data analysis. Hypothesis four and five examined whether gender composition of pairs has an effect on user's enjoyment and engagement depending on different interface styles. The study furthers its explorations through the analysis of covariance of subjects' individual differences. The effects of demographic variables on subjects' enjoyment and engagement based on a pair condition were analyzed.

In this study, data was collected in a number of ways. I based the majority of the results on a statistical analysis of questionnaire responses (enjoyment), the time logs and counts of play time (engagement), and responses to the two collaborative-related questions (collaboration). A qualitative analysis of the observational notes and audio records (collaboration) were used to contextualize the quantitative findings. This chapter describes the general participants' profiles (section 5.2), test results of hypothesis one – enjoyment (section 5.3), hypothesis two – engagement (section 5.4), hypothesis three – collaboration (section 5.5), hypothesis four (section 5.6), hypothesis five (section 5.7), analysis on preference (section 5.8), observations on features (section 5.9) and the exploration of covariate test and analysis of the demographic variables (section 5.10).

5.2 Participant Profile

This study recruited in total 132 children, including 69 boys and 63 girls. Pairings were 23 pairs of boy and boy groups, 20 pairs of girl and girl groups and 23 pairs of boy and girl groups. Twenty two pairs of children worked on cardboard (traditional) puzzles, twenty one pairs worked on desktop puzzles, and twenty three pairs worked on tangible puzzles (Table 5.1).

	Boy-Boy	Boy-Girl	Girl-Girl	Total
PUI	8	7	7	22
GUI	5	7	9	21
TUI	10	9	4	23
Total	23	23	20	66

Table 5.1 Description of participants' pair condition

All child participants were older than 6 but younger than 10 years old. 84% of the children were aged from 7 to 9 years old. 11% were younger than 7 and 5% were older than 9 (Table 5.2). 87% of the children were fluent in English, and the other 23% took English as a second language. 90% of all participants had played a jigsaw puzzle before, and all participants knew how to solve jigsaw puzzles. 64% of the children indicated they really liked to play jigsaw puzzles; 24% indicated that they somewhat liked to play them, and the other 12% indicated that they didn't like to play jigsaw puzzles. 93% participants reported that they had used personal computers, and 92% of them considered themselves good mouse users. 81% of children used the computer at least a few times a week at their home or school, and 64% indicated that used the computer everyday. None of the children had solved the puzzles used in this study.

	Age < 7	Age 7-9	Age >9
Count	14	111	7
%	11%	84%	5%

Table 5.2 Descriptive analysis of participants' age

5.3 Hypothesis One: Enjoyment

As discussed previously, subjects' enjoyment rating on different interface styles is one of the predominant measurements in this research study. Subject's self-reported ratings of enjoyment on the four IMI subscales, which are *Interest and Enjoyment*, *Perceived Competence*, *Perceived Choice*, and *Pressure and Tension*, were used for this measure. This section presents the findings of how interface style affected subjects' enjoyment on different interface styles. The findings from the analysis of the correlations between users' enjoyment and one of the three subscales are presented.

5.3.1 Prediction

The first hypothesis examined how subjects' rating of enjoyment on puzzle tasks was affected by the interface style implemented. It was predicted that user's enjoyment level on the TUI implementation would be higher than that on the GUI or PUI implementations.

5.3.2 Analysis Summary

Preliminary analysis showed that the main data collected from the Likert-style scales were not normally distributed (as discussed in section 4.7). Therefore, non-parametric test tools were required rather than ANOVA tests. A between-subject Kruskal-Wallis test based on the scores of four subscales was conducted in order to examine how subject's *Interest and Enjoyment*, *Perceived Competence*, *Perceived Choice* and *Pressure and Tension* were affected by interface style. The bivariate relation between the *Interest and Enjoyment* subscale and one of the other three subscales can be examined using Spearman's rank correlation.

5.3.3 Main Effect

Descriptive statistics by interface style are show in Table 5.3. The Kruskal-Wallis test indicated that interface style didn't have a significant effect on users' subjective rating of *Interest and Enjoyment*. Statistical analysis on the other three relevant subscales didn't result in any significant differences across the three interface styles either (Table 5.4). Figure 5.1 illustrates the mean scores of these four subscales depending on three different interface styles. Although the difference is not significant, the sum scored subscales show a slight trend. *Interest and Enjoyment* was similar across the three interface styles. Scores for *Perceived Competence* were slightly highest for the PUI condition and lowest for the TUI condition. Similarly, scores for *Perceived Choice* were slightly highest for the PUI condition and lowest for the TUI condition. Inversely, scores for *Pressure and Tension* were slightly lowest for the PUI condition and highest for the TUI condition. Since no significant difference was observed, no evidence exists to support the hypothesis that subjects' enjoyment level was affected by interface style.

Aggregation across interface styles on the sum scored *Interest and Enjoyment* subscale showed that 51 of the 66 pairs (77%) found the puzzle highly interesting and enjoyable (pair mean >= 4.0) independent of interface style. The distributions of the data for this subscale show a restriction on variance.

		N	Mean	Median	S.D.	S. Error
Interest & Enjoyment	PUI	22	4.25	4.43	.751	.160
	GUI	21	4.26	4.57	.956	.209
	TUI	23	4.32	4.36	.527	.110
Perceived Competence	PUI	22	4.51	4.60	.452	.096
	GUI	21	4.20	4.40	.906	.198
	TUI	23	4.13	4.20	.619	.129
Perceived Choice	PUI	22	4.25	4.50	.684	.146
	GUI	21	3.89	3.96	1.012	.221
	TUI	23	3.87	4.00	.845	.176
Pressure & Tension	PUI	22	1.88	1.75	.540	.115
	GUI	21	1.98	1.60	.808	.176
	TUI	23	2.28	2.20	.695	.145

Table 5.3 Descriptive analysis of four IMI subscales

Test Statistics[a,b]

	Interest & Enjoyment	Perceived Competence	Perceived Choice	Pressure & Tension
Chi-Square	.716	3.746	2.611	3.242
df	2	2	2	2
Asymp. Sig.	.699	.154	.271	.198

a Kruskal Wallis Test b Grouping Variable: Interface Style

Table 5.4 Kruskal Wallis test results of the four IMI subscales

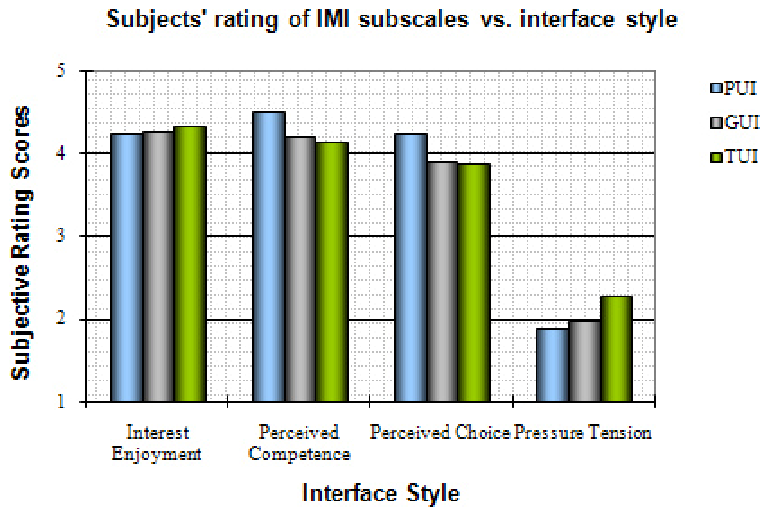

Figure 5.1 The relations of interface style and subjects' rating of four IMI subscales

5.3.4 Correlation Results

I also explored the correlations between *Interest and Enjoyment* and one of the other three IMI subscales. Results of the Spearman's Rho analysis showed that *Interest and Enjoyment* is positively correlated to *Perceived Competence* at an α=0.01 level (two-tailed) with a value of r=0.73. Similarly, *Interest and Enjoyment* is positively correlated to *Perceived Choice* at an α=0.01 level (two-tailed) with a value of r=0.32 (Table 5.5). However, no correlations were found between the *Interest and Enjoyment* and *Pressure and Tension* subscales.

			Interest & Enjoyment	Perceived Competence	Perceived Choice	Pressure & Tension
Spearman's rho	Interest & Enjoyment	Correlation Coefficient	1.000	.730(**)	.320(**)	-.192
		Sig. (2-tailed)	.	.000	.009	.123
	Perceived Competence	Correlation Coefficient	.730(**)	1.000	.339(**)	-.180
		Sig. (2-tailed)	.000	.	.005	.148
	Perceived Choice	Correlation Coefficient	.320(**)	.339(**)	1.000	-.303(*)
		Sig. (2-tailed)	.009	.005	.	.014
	Pressure & Tension	Correlation Coefficient	-.192	-.180	-.303(*)	1.000
		Sig. (2-tailed)	.123	.148	.014	.

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 5.5 Correlation analysis between four IMI subscales

Correlation analyses were also performed depending on different interface styles based on previous significant findings (i.e., *Interest and Enjoyment* and *Perceived Competence*, *Interest and Enjoyment* and *Perceived Choice*). The results shows that *Interest and Enjoyment* are positively correlated to *Perceived Competence* for the PUI condition (at α=0.01 level, r=0.78), GUI condition (at α=0.01 level, r=0.84) and TUI condition (at α=0.01 level, r=0.56) (see Table 5.6). A positive correlation was found between *Interest and Enjoyment* and *Perceived Choice* for the GUI condition (α=0.05 level, r=0.23) and TUI condition (at α=0.05 level, r=0.20), but not for the PUI condition.

	Enj. and Comp.		Enj. and Choice	
	coefficient	Sig.	coefficient	Sig.
PUI	+.781**	.000	+.153	.138
GUI	+.838**	.000	+.226*	.032
TUI	+.560**	.000	+.199*	.048

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 5.6 Correlations across the interface style

5.3.5 Summary

The results show that based on the questionnaire measures, no evidence was found to demonstrate a significant difference among the three interface conditions on the assessed dimensions. However, a weak trend existed in these four IMI subscales. The mean score of the *Interest and Enjoyment* subscale was similar across the three interface styles. Scores for *Perceived Competence* were slightly highest for the PUI condition and lowest for the TUI condition. Scores for *Perceived Choice* were slightly highest for the PUI condition and lowest for the TUI condition. Scores for *Pressure and Tension* were slightly lowest for the PUI condition and highest for the TUI condition. Correlation analysis results indicated that subjects' *Perceived Competence* and *Perceived Choice* are positively correlated to their *Interest and Enjoyment*. However, *Pressure and Tension* is not correlated to *Interest and Enjoyment*.

5.4 Hypothesis Two: Engagement

As discussed previously, subjects' engagement in different interface styles is another important measurement in this research study. Subjects' five time-related measures of engagement, including total play time, time for first completion, time for second completion (if applicable), number of starts, and number of completions were used for this measure.

This section presents the detailed analysis results of how interface style affected subjects' engagement on different puzzle tasks.

5.4.1 Prediction

The second hypothesis examined how subjects' measured times on puzzle tasks were affected by interface style. It was predicted that the user's engagement level on the TUI implementation would be higher than that on the GUI or PUI implementations.

5.4.2 Analysis

Preliminary analysis indicated that the time-log data recorded using minutes and seconds was not normally distributed (as discussed in section 4.7). Therefore, non-parametric test tools were used. A between-subject Kruskal-Wallis test based on the time-log data was conducted in order to examine how subject's overall on-task time, time for first completion, and time for second completion (if applicable) was affected by interface style.

5.4.3 Descriptive Statistics

Descriptive statistics of play time data are shown in Table 5.7. The time-log data based on the 66 pairs of participants revealed that on average, total play time was longest for the GUI condition (13:12) and one minute less for the TUI (11:31) and PUI (10:32) conditions. 48% of the GUI players did not complete the puzzle even once within the total given time (i.e., 15 minutes). Four out of 23 pairs of the TUI players could not finish the puzzle at least once. But only one pair of the PUI players could not finish the puzzle at least once (Table 5.8). For all those pairs who didn't finish their puzzle at least once within the total allocated time (i.e., 15 minutes), two pairs of GUI players quit their play before the 15 minutes end time; however, those TUI and PUI players continued their play and fulfilled the total 15 minutes.

		Total Play Time	Ave. Time to 1st Completion	Ave. Time to 2nd Completion
PUI	Mean	12:24	10:32	6:08
	N	22	22	1
	SD	2.540	2.424	N/A
GUI	Mean	13:17	13:12	N/A
	N	21	21	0
	SD	2.242	2.207	N/A
TUI	Mean	12:22	11:31	6:21
	N	23	23	2
	SD	2.341	3.045	0.14

Table 5.7 Descriptive statistics of task time vs. interface style

	Interface Style											
	PUI				GUI				TUI			
	# of starts		# of completions		# of starts		# of completions		# of starts		# of completions	
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
0	N/A	N/A	1	4.5%	N/A	N/A	10	47.6%	N/A	N/A	4	17.4%
1	14	63.6%	20	90.9%	20	95.2%	11	52.4%	20	87.0%	17	73.9%
2	8	36.4%	1	4.5%	1	4.8%	N/A	N/A	3	13.0%	2	8.7%
Total	22	100%	22	100%	21	100%	21	100%	23	100%	23	100%

Table 5.8 Counts of starts and completions vs. interface style

5.4.4 Main Effect

The relationship between the interface style and the time data was analyzed using a Kruskal-Wallis test. Data including total play time, time for first completion and subsequent play time were examined. As discussed previously (section 3.2), subsequent play time was operationally defined as the time margin between the total play time and the time to its first completion ($t_{subsequent} = t_{total} - t_{1st\ comp}$).

Kruskal-Wallis test results for first completion time showed a statistically significant main effect at the $p<0.005$ level ($\chi^2(2)=11.50$; $p=0.003$). Results for subsequent play time also showed a significant difference at the $p<0.05$ level ($\chi^2(2)=7.60$; $p=0.022$) (Table 5.9). However, results for the total play time didn't show any significant difference depending on interface styles. Figure 5.2 graphically depicts the relative amount of time pairs spent on the first and subsequent play across interfaces.

Test Statistics[a,b]

	Total Play Time	1st completion	Subsequent Play Time
Chi-Square	2.850	11.504	7.600
df	2	2	2
Asymp. Sig.	.240	.003	.022

a Kruskal Wallis Test b Grouping Variable: Interface Style

Table 5.9 Kruskal Wallis test results for time data

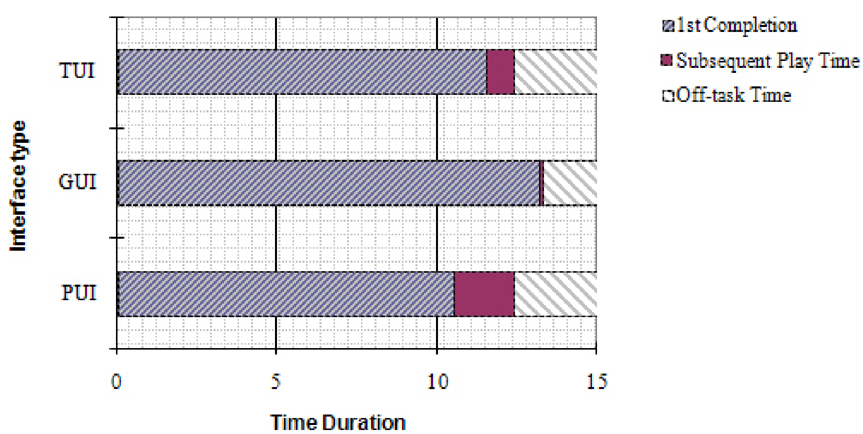

Figure 5.2 Time measures of engagement vs. interface style

Mann-Whitney tests were performed on the previously significant findings (i.e., time to first completion, time to subsequent play). Results indicated that the average time spent on first puzzle completion was significantly longer for the GUI condition than that for the TUI condition at the $p<0.05$ level, Mann-Whitney U=156.5, $p=0.043$, (Table 5.10). Average time spent on first puzzle completion was also significantly longer for the GUI condition than that for the PUI condition at the $p<0.005$ level, Mann-Whitney U=90.5, $p=0.001$, (Table 5.11). However, no significant difference of time to first completion was found between the PUI and TUI condition. Results also indicated that the mean of subsequent play time is significantly longer for the PUI condition than that for the GUI condition at the $p<0.01$ level, Mann-Whitney U=155.0, $p=0.009$, (Table 5.11). The subsequent play time for the TUI condition is also longer than that for the GUI condition but not significantly so. No significant difference of subsequent play time exists between the PUI and TUI condition either.

	1st completion	**Subsequent play time**
Mann-Whitney U	156.500	220.000
Wilcoxon W	432.500	451.000
Z	-2.024	-1.013
Asymp. Sig. (2-tailed)	.043	.311

a Grouping Variable: Interface Style

Table 5.10 Mann-Whitney test result on time-related data (TUI vs. GUI)

	1st completion	**Subsequent play time**
Mann-Whitney U	90.500	155.000
Wilcoxon W	343.500	386.000
Z	-3.429	-2.596
Asymp. Sig. (2-tailed)	.001	.009

a Grouping Variable: Interface Style

Table 5.11 Mann-Whitney test result on time-related data (PUI vs. GUI)

The frequency count of the number of starts indicated that 36% of PUI users had repeat play after their first completion, 13% of TUI users had repeat play, but only 4% of GUI users had repeat play (Table 5.8). The frequency count of the number of completions indicated that 91% of the PUI pairs completed the puzzle once, and 5% of the PUI pairs completed the puzzle twice. Seventy four percent of the TUI pairs completed the puzzle once, and 9% of the TUI pairs complete the puzzle twice. Fifty two percent of the GUI pairs completed the puzzle once, but none of the GUI pairs completed the puzzle twice.

The relationship between the interface style and the number of starts and the number of completions were also analyzed using Kruskal-Wallis tests. The number of times they started over (i.e., repeat play) was significantly different across the three interfaces at the $p<0.05$ level ($\chi^2(2)=7.72$; $p=0.021$) (Table 5.12). Results of the Mann-Whitney tests indicated that repeat play was significantly higher for the PUI than GUI (at the $p<0.05$ level, Mann-Whitney U= 158.0, $p=0.012$) (Table 5.13). The number of repeat plays on TUI was also more than that on GUI but not significantly so. No significant difference of the number of repeat plays was found between TUI and PUI conditions. Results of the Kruskal-Wallis tests indicated that the number of completion was also significantly different across the three interface styles at the $p<0.005$ level ($\chi^2(2)=11.67$; $p=0.003$) (Table 5.12). Mann-Whitney tests results revealed that number of completions was significantly higher for the PUI than GUI at the $p<0.005$ level (Mann-Whitney U=126.0, $p=0.001$) (Table 5.13). The number of completion was also significantly higher for the TUI than GUI at the $p<0.05$ level (Mann-Whitney U=157.5, $p=0.019$) (Table 5.14). However, no significant difference of the number of completion was found between TUI and PUI conditions.

	# of times of starts	# of times of completion
Chi-Square	7.721	11.667
df	2	2
Asymp. Sig.	.021	.003

a Kruskal Wallis Test b Grouping Variable: Interface Style

Table 5.12 Kruskal Wallis test results for number of times of starts and completions

	# of times of starts	# of times of completion
Mann-Whitney U	158.000	126.000
Wilcoxon W	389.000	357.000
Z	-2.516	-3.269
Asymp. Sig. (2-tailed)	.012	.001

a Grouping Variable: Interface Style

Table 5.13 Mann-Whitney test result on the number of starts and completions (PUI vs. GUI)

	# of times of starts	# of times of completion
Mann-Whitney U	221.500	157.500
Wilcoxon W	452.500	388.500
Z	-.944	-2.342
Asymp. Sig. (2-tailed)	.345	.019

a Grouping Variable: Interface Style

Table 5.14 Mann-Whitney test result on the number of starts and completions (TUI vs. GUI)

5.4.5 Summary

The analysis based on subjects' five time-related measures of engagement revealed that results for the time to pairs' first completion had a significant difference across the three interface conditions. Users took significantly longer solving GUI puzzles than solving PUI or TUI puzzles. Results for the users' subsequent play time had a significant difference across the three interface conditions as well. User stayed significantly longer on the PUI condition for their subsequent plays than on the GUI condition. Subsequent play time results of the TUI condition were shorter than that on the PUI and longer than that on the GUI, but no significant difference was found. Results of the number of starts and completions were also significantly different across the three interfaces. The number of repeat plays was significantly higher for the PUI than that for the GUI. Repeat play for the TUI condition is less than that for the PUI and more than that for the GUI, but no significant differences were observed in between these conditions. The results of the number of completions were similar for PUI and TUI tasks, and the results for both conditions were significantly higher than that for the GUI condition.

5.5 Hypothesis Three: Collaborative Interaction

As discussed previously, subject's collaborative interaction on different interface styles was examined because it related to enjoyment and engagement in this study. Both quantitative and qualitative analyses are applied. Statistical analysis is based on subjects' answers to the two collaboration preference questions, whereas qualitative findings are based on informal observations taken during the study. Observation mainly focused on pairs of children's collaboration styles (i.e., independent parallel play, sequential turn-taking play or other styles).

5.5.1 Prediction

The third hypothesis examined how subjects' collaboration style was affected by interface style in solving a spatial puzzle task. It was predicted that users would use a similar collaboration style on the TUI and PUI conditions, but use a different collaboration style on GUI conditions.

5.5.2 Analysis Summary

Collaboration is analyzed using both quantitative and qualitative data. Quantitative analysis is based on subjects' answers to the two collaboration preference questions listed in the post-questionnaire (Appendix C). Qualitative findings are also concluded from informal observations. Preliminary analysis indicated that response data was not normally distributed (as discussed in section 4.7). Therefore, non-parametric tests were used. A between-subject Kruskal-Wallis test was conducted in order to examine how subject's preference of collaboration was affected by interface style.

5.5.3 Quantitative Analysis

The two relevant questions analyzed here are:

Q1: *I would like to do another puzzle like this one with one of my good friends.*

Q2: *I would like to do a similar puzzle like this by myself next time.*

Descriptive statistics results for children's responses to these two questions depending on interface style are show in Table 5.15. For data analysis, "Not at all true" was coded as one, and "very true" was coded as five. A Kruskal-Wallis test result for the first question indicated that interface style didn't have a significant effect on users' preference of doing puzzles with friends. Kruskal-Wallis test results for the second question showed a statistically significant main effect at the $p<0.05$ level ($\chi^2(2)=9.615$; $p=0.008$)(Table 5.16).

		Prefer do again with my friend	Prefer do again by myself
PUI	Mean	4.32	3.21
	N	22	22
	SD	0.783	1.438
GUI	Mean	4.23	2.87
	N	21	21
	SD	1.168	1.470
TUI	Mean	4.15	3.52
	N	23	23
	SD	1.135	1.244

Table 5.15 Descriptive Statistics for the two collaboration related questions

Test Statistics[a,b]

	Prefer do again with my friend	Prefer do again by myself
Chi-Square	1.848	9.615
df	2	2
Asymp. Sig.	.397	.008

a Kruskal Wallis Test b Grouping Variable: Interface Style

Table 5.16 Kruskal Wallis test results for collaboration data

Mann-Whitney tests were performed on the significant findings (i.e., solving puzzle by myself) to examine the details of the differences amongst the three interfaces. The analysis result indicated that children significantly preferred to do the TUI puzzles again by themselves rather than do the GUI puzzles again by themselves at a $p<0.05$ level (Mann-Whitney U=3332.5, $p=0.003$). However, no significant difference was found for either the PUI versus GUI condition or TUI versus PUI condition.

5.5.4 Collaboration Style

The design of the study facilitated observation of different interface styles. I observed that children used a similar collaboration style in both the TUI and PUI conditions, but used a different collaboration style on the GUI condition. Over two-thirds of the pairs solved the TUI and PUI puzzles using independent parallel play in which they seemed to be absorbed in their own activity but they still observed each other's actions and expressions and often copied them (Figure 5.3). For example, each child in the pair often concentrated on a different area of the puzzle. In some cases, their verbalizations revealed a conscious strategy to work cooperatively by dividing puzzles areas between them. "You do the top part and I'll do the bottom." Verbalization in parallel play also often concerned advising the other child where a piece should go. In some cases, children took a directive role where they gave verbal instructions to the other child. This often happened in a pair with an age difference (e.g., one child was 9 years old and the other was 7 year old).

Those pairs using the GUI system used a different strategy to solve puzzles. Despite the single mouse on the GUI puzzle, most pairs found a way to collaborate with each other. About two thirds of the pairs took sequential turns during their play (Figure 5.4). It was also commonly observed in the GUI condition that one child took a dominant or directive role through the whole session. In these cases, the other child often found other ways to collaborate, such as pointing at the screen or giving verbal suggestions to his/her partner.

Figure 5.3 Parallel independent play strategy is commonly observed under TUI & PUI conditions

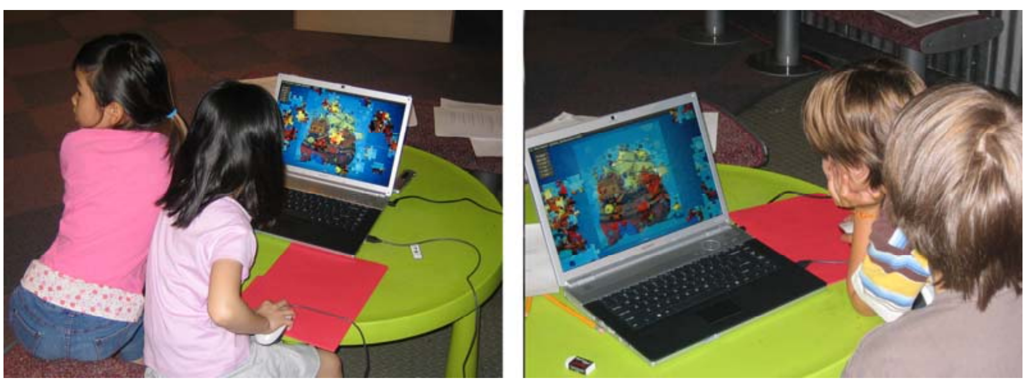

Figure 5.4 Sequential turn taking strategy is commonly used in GUI condition

Through informal observation I observed some cases of other collaboration styles in the three interface conditions. I observed some cases of directive play, in which one child took a directive role to complete the puzzle task. More cases of directive play were observed in the GUI condition, but seldom observed in PUI or TUI conditions. I also observed some cases of cooperative play as mentioned previously. Sometimes, children's verbalizations revealed a conscious strategy to work cooperatively. This strategy was commonly observed in any of the three interface conditions. However, I didn't observe any competitive play strategy or other kinds of collaboration styles in any of the three interface conditions.

5.5.5 Summary

The analysis of collaboration is based on both the quantitative data and the qualitative informal observations. Statistical analysis results indicated that children significantly prefer independent play for solving TUI puzzles rather than solving GUI puzzles. However, no significant difference was found on PUI versus GUI or PUI versus TUI conditions. Results also revealed that no significant difference of children's preferences for collaborating with partners exists across the three interface styles. Observational findings concluded that pairs of children commonly do independent parallel play for solving TUI or PUI puzzles, but do sequential turn-taking play for solving GUI puzzles. Informal observations also indicated that more children adopted a directive play strategy in the GUI condition. Sometimes, children will do cooperative play, however, the use of this strategy is not determined by interface style.

5.6 Hypothesis Four: Gender Composition of Pairs on Enjoyment

As discussed previously, it was proposed that the subject's enjoyment might be affected by gender composition of pairs on different interfaces. In this study, the gender composition of pairs was tested as a second independent variable for measuring subject's enjoyment. This section presents a detailed analysis results of how gender composition of pairs affects subjects' enjoyment differently depending on interface style.

5.6.1 Prediction

The fourth hypothesis examines how gender composition of pairs affects subjects' enjoyment differently depending on interface. It was predicted that the interaction of interface style and gender composition of pairs has a significant effect on children's enjoyment on solving a spatial puzzle task.

5.6.2 Analysis Summary

Preliminary analysis indicated that the main data collected from Likert-style scales were not normally distributed (as discussed in section 4.7). Therefore, non-parametric test tools were used instead of using traditional two-way ANOVA test. A Friedman's test was performed on the scores of four IMI subscales to examine how subject's *Interest and Enjoyment*, *Perceived Competence*, *Perceived Choice* and *Pressure and Tension* were affected by gender composition of pairs depending on interface style.

5.6.3 Results

Descriptive statistics results of the four IMI subscales under different gender pairing conditions depending on different interface styles are shown in the following four tables (Table 5.17, Table 5.18, Table 5.19, and Table 5.20).

Interest & Enjoyment		PUI	GUI	TUI
Boy & Boy	Mean	4.44	4.69	4.29
	N	8	5	10
	SD	.392	.374	.733
Boy & Girl	Mean	3.63	4.18	4.19
	N	7	7	9
	SD	1.22	.952	.827
Girl & Girl	Mean	4.65	4.07	4.04
	N	7	9	4
	SD	.560	1.16	1.06

Table 5.17 Descriptive Statistics for *Interest & Enjoyment* vs. interface style & gender pairing

Perceived Competence		PUI	GUI	TUI
Boy & Boy	Mean	4.54	4.72	4.02
	N	8	5	10
	SD	.610	.681	.894
Boy & Girl	Mean	4.33	4.09	4.34
	N	7	7	9
	SD	.530	.833	.565
Girl & Girl	Mean	4.66	3.99	3.60
	N	7	9	4
	SD	.486	1.17	.821

Table 5.18 Descriptive Statistics for *Perceived Competence* vs. interface style & gender pairing

Perceived Choice		PUI	GUI	TUI
Boy & Boy	Mean	4.46	4.60	3.63
	N	8	5	10
	SD	.668	.499	.914
Boy & Girl	Mean	4.20	3.70	3.93
	N	7	7	9
	SD	1.14	.950	.857
Girl & Girl	Mean	4.07	3.64	3.83
	N	7	9	4
	SD	.664	1.23	1.27

Table 5.19 Descriptive Statistics for *Perceived Choice* vs. interface style & gender pairing

Pressure & Tension *		PUI	GUI	TUI
Boy & Boy	Mean	2.10	1.68	2.20
	N	8	5	10
	SD	.700	.900	.894
Boy & Girl	Mean	1.57	2.16	2.22
	N	7	7	9
	SD	.367	.842	.912
Girl & Girl	Mean	1.94	2.00	2.40
	N	7	9	4
	SD	.630	.823	.786

* represents the scores were interpreted reversely

Table 5.20 Descriptive Statistics for *Pressure and Tension* vs. interface style & gender pairing

The relationship among the interface style, gender pairing and subject's enjoyment was analyzed using a Friedman's test. Friedman's test results for the subscale of *Interest and Enjoyment* showed a statistically significant main effect at the $p<0.001$ level ($\chi^2(2)=411.8$; $p=0.000$). Results for the subscales of *Perceived Competence* and *Perceived Choice* also showed significant differences at the $p<0.001$ level ($\chi^2(2)=453.03$; $p=0.000$; $\chi^2(2)=390.85$; $p=0.000$).

However, no significant difference was found on the test results of the *Pressure and Tension* subscale.

Figure 5.5 graphically depicts the *Interest and Enjoyment* subscales ratings by different gender compositions of pairs on three different interface styles. While all gender pairings' mean scores on the *Interest and Enjoyment* subscale are nearly the same in the TUI condition, boy-boy pairs have a higher mean than girl-girl pairs or boy-girl pairs on the GUI condition. However, boy-girl pairs have a lower mean than boy-boy pairs and girl-girl pairs on the PUI condition. Girl-girl pairs' scores are higher than boy-boy pairs or boy-girl pairs on the PUI condition.

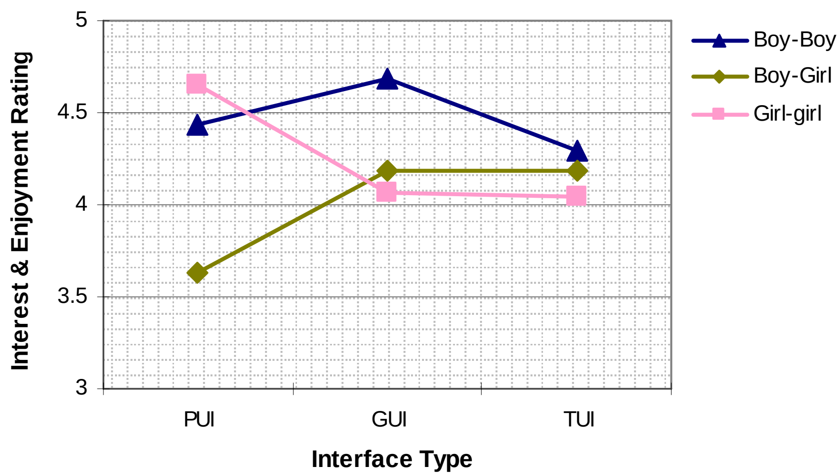

Figure 5.5 Rating of *Interest and Enjoyment* by different gender pairing on three interfaces

Figure 5.6 graphically depicts the ratings of *Perceived Competence* subscales by different gender compositions of pairs on three different interface styles. Girl-girl pairs show a higher mean than boy-boy pairs and boy-girl pairs on the PUI condition. Girl-girl pairs show a lower mean than boy-boy pairs and boy-girl pairs on the TUI condition. Boy-boy pairs report a higher mean than boy-girl pairs or girl-girl pairs on the the GUI condition.

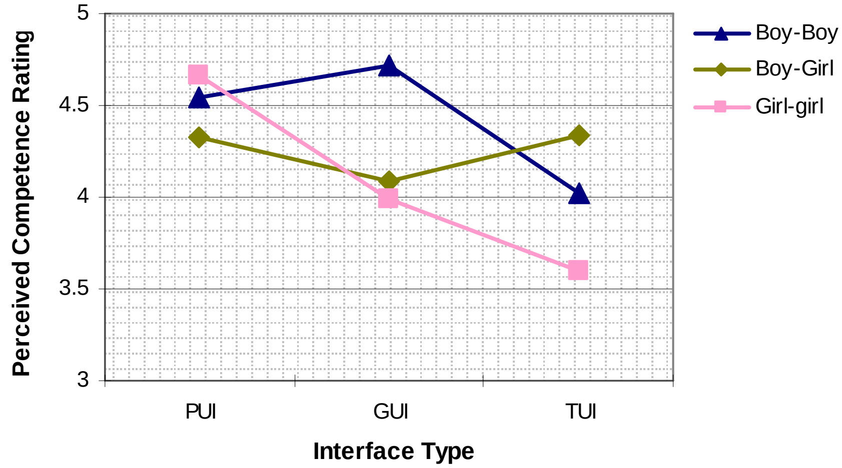

Figure 5.6 Rating of *Perceived Competence*
by different gender pairing on three interfaces

Figure 5.7 graphically depicts the ratings of *Perceived Choice* subscales by different gender compositions of pairs on three different interface styles. While all gender pairings' mean scores on the *Perceived Choice* subscale are nearly the same in the TUI condition, boy-boy pairs report a higher mean on both the PUI and GUI conditions than boy-girl pairs or girl-girl pairs do. Girl-girl pairs and boy-girl pairs report similar means across all three interface conditions.

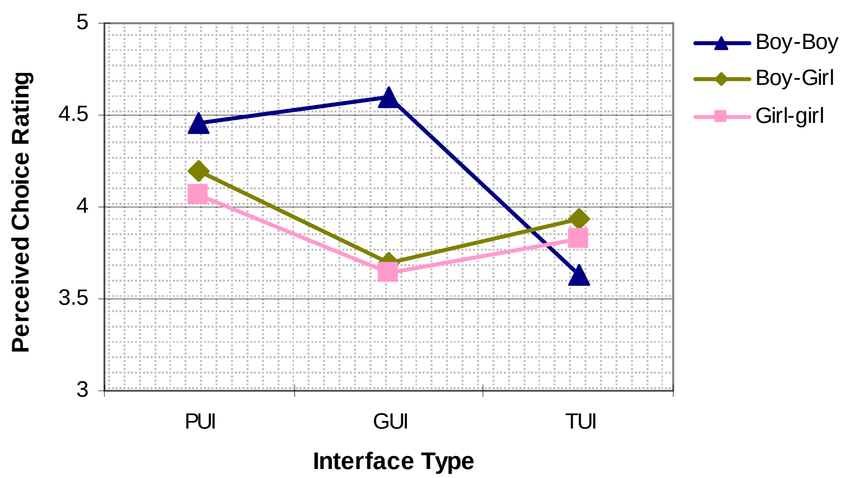

Figure 5.7 Rating of *Perceived Choice*
by different gender pairing on three interfaces

5.6.4 Summary

The results indicate that gender composition of pairs had a significant effect on user's self-reported rating of *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice* subscales across the three different interface styles. However, no significant effect of gender composition of pairs was found in the subject's *Pressure and Tension* subscale on different interface styles.

All gender pairings' mean scores on *Interest and Enjoyment* and *Perceived Choice* subscales are nearly the same in the TUI condition. Boy-boy pairs consistently reported higher means on *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice* subscales than boy-girl pairs or girl-girl pairs did in the GUI condition. Girl-girl pairs consistently reported higher means on *Interest and Enjoyment* and *Perceived Competence* in the PUI condition than the boy-boy pairs and boy-girl pairs. One interpretation is that boy-boy pairs enjoy desktop puzzles more, whereas girl-girl pairs enjoy traditional cardboard puzzles more.

5.7 Hypothesis Five: Gender Composition of Pairs on Engagement

As discussed previously, it was proposed that the subject's engagement might be affected by gender composition of pairs on different interface styles. In this study, gender composition of pairs was tested as a second independent variable for measuring subjects' engagement. This section presents the detailed analysis results of how gender composition of pairs affects subjects' engagement differently depending on interface style.

5.7.1 Prediction

The fifth hypothesis examines how gender composition of pairs affects subjects' engagement differently depending on interface style. It was predicted that the interaction of interface style and gender composition of pairs has a significant effect on children's engagement.

5.7.2 Analysis

Preliminary analysis indicated that the time-log data recorded using minutes and seconds was not normally distributed (as discussed in section 4.7). Therefore, non-parametric test tools were used instead of using a traditional two-way ANOVA test. A Friedman's test was performed on the time-log data in order to examine how subject's overall on-task time, time to first completion, and time to subsequent play was affected by the gender composition of pairs on different interfaces.

5.7.3 Results

Descriptive statistics results of pair's time to first completion and time to subsequent play by different gender compositions of pairs on different interface styles are shown in the following two tables (Table 5.21 and Table 5.22).

Time to 1st Comp.		PUI	GUI	TUI
Boy & Boy	Mean	11:32	13:51	11:38
	N	8	5	10
	SD	2.08	1.57	3.08
Boy & Girl	Mean	10:03	13:07	11:04
	N	7	7	9
	SD	2.65	1.15	3.46
Girl & Girl	Mean	9:53	12:55	12:09
	N	7	9	4
	SD	2.53	3.10	2.54

Table 5.21 Descriptive Statistics for the time to 1st Completion vs. interface style & gender pairing

Subsequent play time		PUI	GUI	TUI
Boy & Boy	Mean	0:30	0	0:39
	N	8	5	10
	SD	1.41	N/A	2.04
Boy & Girl	Mean	1:10	0:14	1:29
	N	7	7	9
	SD	2.55	.612	2.94
Girl & Girl	Mean	4:05	0	0
	N	7	9	4
	SD	3.19	N/A	N/A

Table 5.22 Descriptive Statistics for the time to subsequent play vs. interface style & gender pairing

The relationships among the interface style, gender composition of pairs and subject's time-related data were analyzed using Friedman's tests. Friedman's test results for the time to first completion shows a statistically significant effect at the $p<0.001$ level ($\chi^2(2)=109.8$; $p=0.000$). Results for the time to subsequent play also show a significant effect at the $p<0.001$ level ($\chi^2(2)=55.70$; $p=0.000$). Figure 5.8 graphically depicts the relations of user's first completion time, subsequent play time and off-task time by different gender compositions of pairs on different interface styles.

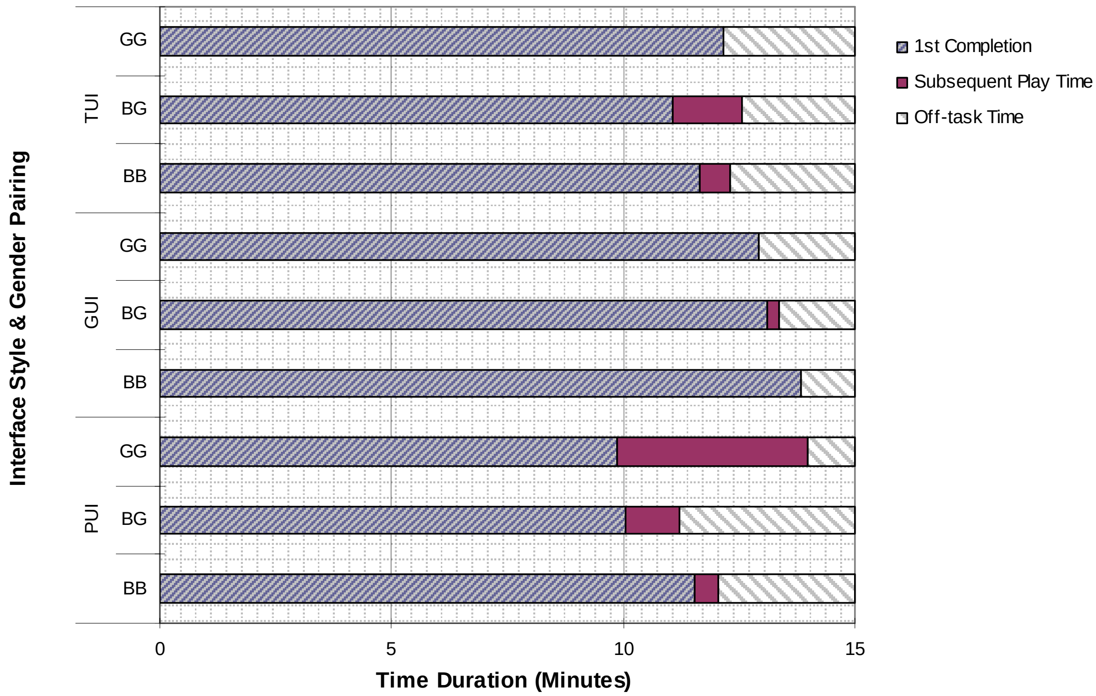

Figure 5.8 Subject's time-related measures of engagement vs. interface style and gender pairing

While all gender pairings spent longest for their first completions in the GUI condition, but spent shortest for their first completions in the PUI condition. Boy-boy pairs spent longer on their first completion than girl-girl pairs did on either the PUI or GUI condition. None of the boy-boy pairs or girl-girl pairs stayed for subsequent plays in the GUI condition. None of the girl-girl pairs stayed for subsequent plays in the TUI condition. However, all gender pairings reported subsequent play in the PUI condition. Girl-girl pairs reported longer time for their subsequent play in the PUI condition.

5.7.4 Summary

Statistical results indicated that gender composition of pairs has a significant effect on subjects' time to first completion and time to subsequent play on three different interfaces.

Overall, all gender pairings spent longest for first completions in the GUI condition, but spent shortest for first completions in the PUI condition. Girl-girl pairs spent shorter time on their first completions than boy-boy or boy-girl pairs did on both GUI and PUI conditions. All gender pairings had subsequent play on PUI condition. Girl-girl pairs stayed longer for their subsequent play in the PUI condition than boy-boy pairs or boy-girl pairs did.

5.8 Preference

Subjects' comments of their preference on the three different puzzle implementations were predominantly used to contextualize the study findings. Data were collected through digitized audio recordings of the two open questions. These two questions were:

Q1: *Tell me what things you most like about doing this jigsaw puzzle.*

Q2: *Tell me what things you most didn't like about doing this jigsaw puzzle.*

Thematic analysis of the preference comments resulted in the identification of three other common themes: challenge versus task achievability; reference picture assistance; and help through collaboration.

5.8.1 Results

Many participants commented that the puzzles were fun and enjoyable. They also liked the illustration style and the themes of the puzzle. About 40% of the participants commented that nothing made them dislike the puzzle they played.

Children commented that the puzzle was challenging but that they liked it because they could finish it within the allocated length of time. Some children commented that they were concerned about how much time they had already spent and how much time they still had left for solving the puzzle in the progress of play. This finding is in line with guidelines proposed by Salen and Zimmerman (2003), which state that an enjoyable game balances challenge against possibility of winning. Malone (1980) also proposed that challenge is one essential characteristic of enjoyable computer games. Children also commented that they liked getting help during play from either the reference pictures or their partner (collaboration). This result was consistent with our observational finding on their collaboration and use of the reference picture.

Some children indicated that they did not like the circumstance when the picture underlying the puzzle was turned off (perhaps by their partner). A few children mentioned that they disliked feeling pressured due to the time limitation. This comment was more frequently collected from the pairs in the GUI condition. Some children complained that there were too many pieces in GUI puzzles (which actually had fewer pieces than the TUI or PUI puzzles had).

5.8.2 Summary

The preference analysis based on subjects' answers to the two open-ended questions did not reveal big difference across the three interfaces. Overall, children commented on all the puzzles as fun and enjoyable. Subjects commonly indicated that they liked getting help while solving puzzles either from reference images or from their partner. Some users indicated that they disliked feeling pressured due to the time limitation, especially in GUI condition.

5.9 Observations on Features

As discussed previously, through the informal observation, I also looked at whether the important features of a tangible interface has an impact on children's enjoyment or engagement in solving a spatial puzzle task. These features include direct physical manipulation, integration of input and output space (i.e., reference images), and digital auditory and visual feedbacks.

5.9.1 Direct Physical Manipulation

I expected differences in the physical manipulation of puzzle pieces due to the differences in interaction styles between the TUI and PUI (direct interaction) and GUI (indirect interaction). However, I did not observe big differences in the number or pattern of movements, rotations or connections across the three conditions.

I observed that some young children had difficulty rotating GUI based pieces since this action required the child to simultaneously hold down one button while clicking the other. I observed that over half of the boys preferred using a touch-pad rather than using the mouse on the GUI system, effectively moving to a more direct style of interaction. I also observed that children were much more active in terms of body movement in both the PUI and TUI conditions. For example, some children moved themselves around the table rather than moving the puzzle pieces. Some children made the puzzle in an upside-down direction (Figure 5.9). This form of perspective taking was not possible in the single access GUI condition.

Figure 5.9 Children made puzzles in an upside-down direction

5.9.2 Integration of Input and Output Space

In all three conditions I provided the option to build the puzzle "on top" of the puzzle picture. In PUI and TUI conditions, this meant that the reference picture occupied an integrated input and output space (i.e., puzzle pieces lay on top of display). In the GUI condition the picture was displayed in the output space (i.e., the screen) separated from the input space (i.e., the mouse). I observed that most pairs in all conditions built the puzzle on top of the image during the first attempt to solve the puzzle. Through my observation, it was evident that children preferred image matching rather than other kinds of matching (e.g., color or shape matching). However, no perceivable benefit was observed to having the input and output spaces integrated in space.

5.9.3 Digital Auditory and Visual Feedback

As discussed previously, both the GUI and TUI puzzles provide their users digital auditory and visual feedback, but not for the PUI puzzles. I observed that some children could quickly recognize this feature at the beginning of their playing and use its help for solving puzzles, but other children just ignored it and solved the puzzles by themselves through the whole play period. In an analysis of the chronology of play sessions, no perceivable benefit was observed to having the digital auditory and visual feedback.

5.9.4 Summary

The informal observations over the features revealed some findings. In terms of users' direct physical manipulations, no difference was observed. Observations on the integration of input and output space indicated that most pairs in all conditions built the puzzle with the help of reference images. However, no perceivable benefit was observed for using different integrated representations. Observations on digital auditory and visual feedback indicated that no perceivable benefit was observed for having this feature either.

5.10 Demographic Variables

Demographic data was collected and analyzed to ensure that test subjects formed a relatively homogeneous group and to determine if any of the demographic variables affect children's enjoyment or engagement. In order to investigate the latent variables and increase the precision of comparisons between groups, an analysis of covariance was performed on several important prognostic variables. The analyses of covariance were performed on both the enjoyment and engagement data to examine whether they were affected by other factors.

Other than the gender composition of pairs, the sample population formed a relatively homogeneous group in terms of other demographic variables including age, native language, computer experience, competence of mouse control, puzzle experience, preference of puzzle, and preference of theme.

5.10.1 Age Group

As discussed previously, age group was characterized into four groups according to the recruiting (Section 4.7). Descriptive analysis results of age group on the main enjoyment and engagement data are shown in Table 5.23 and Table 5.24. Results indicated that older age composition of pairs spent shorter time on their first completion than the younger age composition of pairs. However, older age composition of pairs reported less enjoyment than the younger age composition of pairs. Statistical covariate analysis didn't report any significant effect by age group.

Age Group	N	Mean Enjoyment	Mean Competence	Mean Choice	Mean Pressure
Group 1	1	4.29	3.80	4.30	1.40
Group 2	12	4.33	4.28	3.67	2.01
Group 3	47	4.23	4.28	4.10	2.02
Group 4	6	4.06	4.08	3.58	2.33

Table 5.23 Means of IMI subscales (enjoyment) per age group

Age Group	N	Time to 1st Completion	Time to Subsequent Play
Group 1	1	15:00	0:00
Group 2	12	11:23	1:17
Group 3	47	12:01	0:43
Group 4	6	9:35	2:13

Table 5.24 Means of time-related data (engagement) per age group

5.10.2 Computer Experience

As discussed previously, subject's computer experience was characterized into four groups (Section 4.7). Since the engagement data was collected under pair conditions, it didn't meet individual measures. Thus, this measure was used to examine the enjoyment data only. Descriptive analysis results of the four IMI subscales depending on computer experience are shown in Table 5.25.

Computer Experience	N	Mean Enjoyment	Mean Competence	Mean Choice	Mean Pressure
None	9	3.48	3.98	3.67	1.76
Monthly	16	4.61	4.40	4.02	1.85
Weekly	41	4.21	4.35	3.95	2.02
Daily	66	4.26	4.21	4.02	2.13

Table 5.25 Means of IMI subscales (enjoyment) vs. computer experience

Statistical covariate analysis suggests that computer experience is a strong covariate that could vary user's subjective rating on the *Interest and Enjoyment* subscale with the interaction of interface styles. Friedman's test was performed to test how computer experience affects children's enjoyment on different interface styles. Results for the *Interest and Enjoyment* subscale showed a statistically significant effect at the p<0.001 level ($\chi^2(2)=107.26$; $p=0.000$). Results for the *Perceived Competence* and *Perceived Choice* subscales also showed significant differences at the $p<0.001$ level ($\chi^2(2)=175.03$; $p=0.000$; $\chi^2(2)=151.36$; $p=0.000$). However, no significant difference was found on the result of *Pressure and Tension* subscale depending on subjects' computer experience on different interfaces.

Figure 5.10 graphically depicts mean scores of the *Interest and Enjoyment* subscales by computer experience depending on interface style. Figure 5.11 graphically depicts mean scores of the *Perceived Competence* subscales by computer experience depending on interface style. Figure 5.12 graphically depicts mean scores of the *Perceived Choice* subscales by computer experience depending on interface style. Results indicated that pairs of children who have no experience on computers before reported significant lower scores on all the three subscales in the GUI condition. However, subjects who had some computer experience, no matter how much it was, reported similar ratings for these three subscales on the three different interfaces. I also found that the significant result in the GUI condition (lack of computer experience) comes from only one pair. Gender composition of this pair is girl and girl. I suggest that the result could be in line with the findings I have indicated in Section 5.6 (i.e., gender composition of pairs on enjoyment).

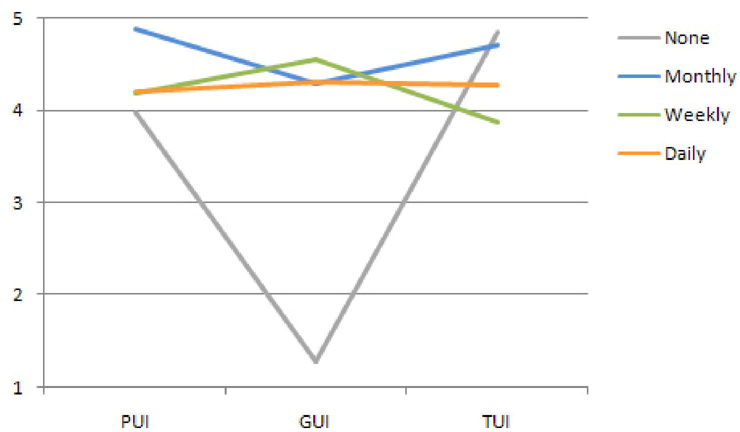

Figure 5.10 ***Interest & Enjoyment*** **subscales vs. interface style & computer experience**

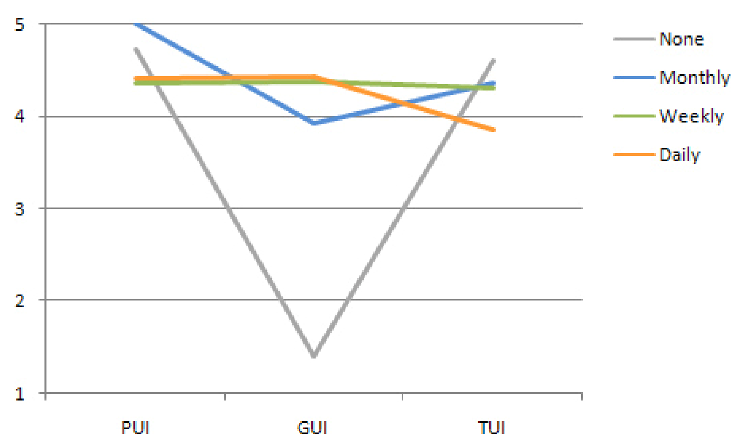

Figure 5.11 ***Perceived Competence*** **subscales vs. interface style & computer experience**

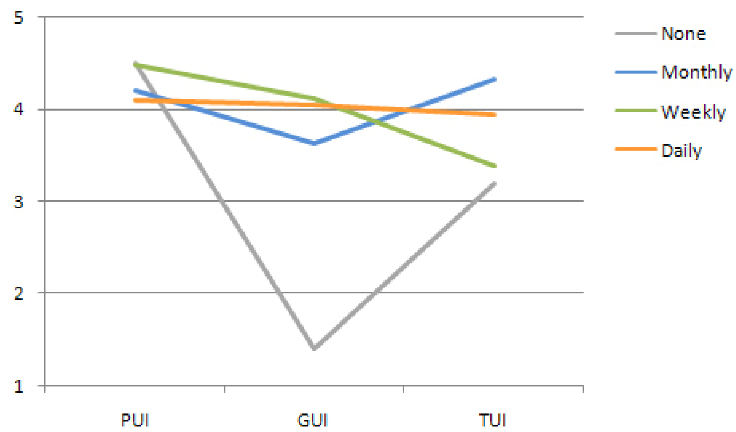

Figure 5.12 ***Perceived Choice*** **subscales vs. interface style & computer experience**

5.10.3 Theme

As discussed previously, the design of this study included two themes for the puzzle images (Section 3.3). Descriptive analyses results of the effect of themes on enjoyment and engagement data are shown in Table 5.26 and Table 5.27. Results showed that the means did not vary significantly between the two themes. Statistical covariate analysis didn't report any significant effect from theme either.

Theme	N	Mean Enjoyment	Mean Competence	Mean Choice	Mean Pressure
Pirate/Boat	36	4.18	4.22	4.16	1.98
Wizard/Castle	30	4.38	4.34	3.82	2.14

Table 5.26 Means of IMI subscales (enjoyment) vs. theme

Theme	N	Time to 1st Completion	Time to Subsequent Play
Pirate/Boat	36	11:58	0:54
Wizard/Castle	30	11:25	0:59

Table 5.27 Means of time-related data (engagement) vs. theme

5.10.4 Others

Similarly, no significant effects were found through the analysis of covariance based on native language, competence of mouse control, puzzle experience, and preference of puzzle.

5.10.5 Summary

The sample population formed a relatively homogeneous group in terms of demographic variables including age, theme, native language, user's competence on mouse control, puzzle experience, and preference of puzzle. However, covariate analysis suggests that computer experience has an effect on children's self-reported *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice*. Pairs of children who were lacking in computer experience reported a significantly lower score on the *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice* subscales for the GUI condition.

In Chapter Six – Discussion, I discuss the context for these findings by addressing the research questions laid out in Chapter Two and discussing the design implications with respect to the design of tangible user interfaces for school aged children. I also compare the results with related research in this field.

CHAPTER 6: DISCUSSION

6.1 Overview

This study is the only known comparative experimental study that compares the use of tangible, graphical and physical (traditional) user interface for evaluating subject's enjoyment and engagement level on a playful learning task by school-aged children (other related studies include Fails et al., 2005; Terrenghi et al., 2006). The three puzzle implementations share many characteristics between each other in order to facilitate a valid comparison (as discussed in Section 3.4). In this chapter, based on the research questions I posed in Chapter Two, I give an overview of the results outlined in Chapter Five related to five hypotheses listed in Chapter Three. I discuss the implications of the evidence from the present study in relation to past research in the field of HCI, physical, social and cognitive theories related with children, and bring forward some of the limitations in this study.

In this chapter, I first discuss the implications and limitations based on the findings of enjoyment in section 6.2. Although statistical results did not reveal a predictive significant difference of enjoyment across the three interfaces, there were some interesting trends and evidence. Most discussions of this topic are related to intrinsic motivation theories. I next discuss the findings of engagement in section 6.3. Engagement analysis results indicated that differences existed across the three interface styles. I discuss the evidence from the study and propose a benefit of physical manipulation according to cognitive reflection theories. In section 6.4, I discuss how interface style affects children's collaborative interactions. In section 6.5, I discuss how gender composition of pairs can affect children's enjoyment and engagement depending on interface style. In section 6.6, I discuss the demographic findings (i.e., computer experience), and how this finding relates to other findings I discussed previously. In section 6.7, I present the design implications resulting from the findings of this research study, and generalizations relating to the design of tangible interfaces for children.

6.2 Discussion: Enjoyment

6.2.1 The IMI Subscales

Many researchers have explored the possibility of using TUI system to augment children's play (e.g., Fails et al., 2005; Raffle et al., 2006). However, this research study produced no evidence to demonstrate that the difference of school-aged children's enjoyment is directly related to interface styles on solving a spatial puzzle task. Pairs of children who interacted with GUI puzzles (predicted as least enjoyable) did not report significantly lower levels of enjoyment than in the other two conditions. The design of this study randomly assigned pairs of children to solve puzzles under different environments. The analysis results of enjoyment suggest that when we separate children into groups for solving a puzzle, no significant difference of enjoyment is found on different interface style. It is possible that peer influence or opportunities to participate may account for their enjoyment levels (Brennan et al., 1997; Carroll et al., 2001). Thus, it maybe worth looking at different ways of grouping other than mean splits (Carroll et al., 2001). For examples, pairs can decide to choose one of the six puzzles and test on it. Alternatively, instead of doing a between-subject study, I can do a within-subject study; however, the interval between each two sessions (i.e., using different interfaces) should be longer enough to avoid the possibility of order effect.

The IMI questionnaire, which used 1-5 Likert style happy to sad face approach, may not be a reliable and valid indication of children's real enjoyment. Children at this age might have difficulty mapping their evaluation of their real enjoyment by using the happy or sad face metaphor. Subjects' inaccurate responses to the questionnaires might cause invalid results.

6.2.2 Correlations between the IMI Subscales

The evidence from the correlation analysis confirmed past research where children's perceived competence was related to enjoyment in physical activities (Brustad, 1993; Carroll et al., 2001). These findings suggest a dynamic relationship might exist among all these two IMI subscales. Dynamic interrelations could be further explored in specific contexts. For example, we can look at the interrelations among interface styles, gender composition of pairs, *Interest and Enjoyment*, *Perceived Competence*, and *Perceived Choice*, or the interrelations among interface styles, children's age, *Perceived Competence*, *Perceived Choice* and *Pressure and Tension* subscales.

The results of this study also reported positive correlations between subject's *Interest and Enjoyment* and *Perceived Choice*. This finding upholds past research findings that if children are allowed to participate in freely chosen activities for the sake of participating in the activity, a task orientation is more likely to emerge; therefore, activities which encourage participants to be task-involved will be more likely to foster feelings of intrinsic motivation (Duda et al., 1995; Mandigo et al., 2000). However, simply allowing choice does not always guarantee increased motivation (Mandigo et al., 2000). Design researchers should take notice for facilitating appropriate scenarios in which choice can be optimized when designing a study for children.

Subjects' verbal debriefs included comments that although the puzzle was challenging they still liked it because they could finish it within the allocated length of time. This finding was in line with the previous research that has demonstrated that when children were optimally challenged while participating in a task, they were more likely to experience enjoyment and be interested in the activity (Danner et al., 1981; Salen et al., 2003). When individuals take part in such activities that challenge them in a positive way, their competence is enhanced. This enhanced competence in turn leads to individuals feeling intrinsically motivated to participate (Mandigo et al., 2000). Understanding the relative importance of game challenge and interface style is worthwhile, as most researchers agree that an enjoyable game balances challenge against the possibility of winning (Salen and Zimmerman, 2003).

6.3 Discussion: Engagement

The statistical results support the prediction that interface style has an effect on user's engagement levels for solving puzzles. During experimental tasks, subject's subsequent play time was significantly longer for solving traditional cardboard puzzles than for solving GUI puzzles. Informal observations on features revealed that young children had some problem with indirect manipulation. These results suggest that there might be a benefit to physical manipulation of concrete objects that can contribute to young user's engagement on a specific task. This partially confirmed the findings from several recent research studies, which suggested that embedding technology into the everyday physical world, rather than simply presenting them with desktop applications, maybe beneficial to young children (Fails et al., 2005). Bruner (1973) proposed that effective learning takes place when meaning is taken from experience with the world, when children through their own experience discover what is "going on in their own heads." Physical engagement with something creates an involvement and activeness in learning (Price et al., 2004).

The evidence of subsequent play time results indicated that children stayed significantly longer for the PUI and TUI conditions than they did for the GUI condition. Danner (1981) indicated that when children are interested in the activity, they will spend longer amounts of time doing the task. This confirmed the viewpoints of many researchers who view engagement as relating to enjoyment (Csikszentmihalyi, 1990; Markopoulos, 2003), which seems appropriate for many of the playful activities of children. However, this study separates the analysis of enjoyment and engagement. The dynamic interrelations between enjoyment and engagement are definitely worth investigating in future study.

The other important finding from this study is related to the difference of repeat play times on the three different interfaces. I counted more repeat plays and completion events on the TUI and PUI conditions than that on the GUI condition. In combination with the advantage of physical manipulation I discussed previously, this finding seems to suggest that direct interaction with the physical manipulation of concrete objects might benefit young users in achieving problem solving tasks, by affecting their level of engagement as well. Alternatively, the finding might also raise the necessity of considering children's cognitive reflection abilities. Brustad (1998) proposed that abstract represents are not developmentally appropriate for young children. Children have limited cognitive capabilities to be able to focus on more than one thing at a time. The separation of physical manipulation and digital representation may interfere in children's cognitive thinking progress on a problem-solving task. This may also result in a lower level of engagement. The operationalization of engagement for children in this research study remains problematic. Task time reflects a mix of engagement and cognitive difficulty in problem solving games. Thus, time to pairs' first completion may more accurately reflect cognitive difficulty rather than engagement. I suggest that time of pairs' subsequent play and number of repeats may be a better indication for engagement in this case. Multiple measures are needed for future studies.

6.4 Discussion: Collaboration

In this study, I observed different collaboration styles used by pairs of children on three different interface styles. Informal observations indicated that children predominantly used independent parallel play for the TUI and PUI conditions, but used sequential turn-taking strategy for the GUI condition. This finding supports the assumption that collaboration style is related to interface style and input design. The multiple access points afforded by a tabletop game (physical (traditional) and tangible) combined with enough space to move supported parallel independent play rather than sequential turn taking.

Quantitative analysis of subjects' responses to the collaboration preference questions indicated that children highly prefer playing puzzles with a partner, rather than by themselves. In an analysis of the chronology of play sessions, I found that all interface styles share the same pattern. Children's verbal debriefing also indicated their preference for solving puzzle under a collaborative condition. This finding supports past research suggesting that a collaborative environment is more likely to elicit increased intrinsic motivation (Inkpen et al., 1995; Sluis et al., 2004).

In the main experiment, our recruiting arbitrarily grouped children into pairs according to their visiting sequence. Some pairs of children were familiar with each other because they are friends or relations and came to the study together. However, there were still many pairs of participants who did not know each other before participating in this study. The study design did not account for this factor. This may be worth future exploration in order to better understand how children's collaborative interaction depends on whether they knew their partners or not.

6.5 Discussion: Gender Composition of Pairs

Significant differences of subject's *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice* on the three interface styles were reported under different gender pairing conditions. Statistical results indicated that the gender composition of pairs has a significant effect on children's enjoyment (i.e., *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice* subscales) and engagement (i.e., time to first completion, time to subsequent play) on the three different interface styles.

Results indicated that boy-boy pairs, in general, reported a significantly higher level of *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice* than girl-girl pairs did in the GUI condition. Girl-girl pairs reported higher means of enjoyment in the PUI condition than boy-boy pairs or boy-girl pairs did. These findings are in line with conclusions from previous research that have identified gender differences in attitudes and use of computers. Siann et al. (1988) found that boys were more positive towards computers and were more confident in using computers than girls. Boy-boy pairs in the present study perceived their competence in puzzle tasks to be higher than the girl-girl pairs in the GUI and TUI conditions. This is in line with past research that showed that gender had an effect on perceived competence (Brustad, 1993; Carroll et al., 2001). Results from this study and past research indicate that gender differences commonly exist in technology embedded interaction design studies, especially for those involving child users.

Due to the limitation of our recruiting, gender composition of pairs includes both same gender and mixed gender compositions in this study. Through the analysis of this study, I found that the mixed gender composition has less contribution to the findings than same gender composition does. I suggest that future study should only involve same gender composition if possible.

6.6 Discussion: Demographic Variables

The sample population formed a relatively homogeneous group in terms of demographic variables including age, theme, native language, user's competence on mouse control, puzzle experience, and preference of puzzle. However, covariate analysis tests suggest that computer experience has an effect on children's self-reported enjoyment (i.e., *Interest and Enjoyment*, *Perceived Competence* and *Perceived Choice* subscales). However, I found only one pair of children lacking in computer experience that participated in our study in the GUI condition. This pair reported extremely low enjoyment on the puzzle task. I suggest the results must be interpreted with caution. Detailed analysis indicated that the gender composition of this pair is girl and girl. This finding indirectly supports the findings I discussed previously (i.e., gender composition of pairs) that girls usually enjoy computer tasks less than boys do.

6.7 Design Implications

Based on the findings and discussions outlined previously in this paper, I present several general implications for the design of tangibles for children.

6.7.1 TUIs are not inherently more enjoyable than other user interface styles

Quantitative analysis indicated that Tangible User Interfaces (TUIs) are not inherently more enjoyable for children in solving a spatial puzzle task than other user interface styles. The finding contradicts the implicit assumption that tangible interfaces are more natural, and thus they are more enjoyable for children. This result also contradicts Fails' (2005) conclusion that tangible interfaces are more interesting than graphical interfaces. This suggests that there is no reason to use tangibles just for creating an enjoyable experience with children. Other reasons might prompt designing tangibles for children.

6.7.2 Children liked interface features that reduce task challenge

Through informal observation, I found that children preferred getting help from reference images while solving puzzles. However, by doing this they changed tasks from form/shape matching to image/colour matching, which is easier. Other than that, direct physical manipulation was also reported as easier and created less frustration for children than indirect interaction (i.e., mouse, touchpad), as discussed previously. These features can reduce task challenge. However, learning sometimes may require challenge. For children this may mean better supporting their tasks, but hopefully not at the expense of learning spatial skills.

6.7.3 Interface design should support task achievability balanced with challenge

Keeping children enjoyed and engaged requires that the UI design (and activity design) support task achievability balanced with challenge. As discussed previously, optimal challenge in a task can likely achieve children's enjoyment and thus makes them interested in the activity (Danner et al., 1981; Salen et al., 2003). I suggest that interface designers should take into account the balance of task achievability and challenge when designing an enjoyable and engaging tangible experience for children.

6.7.4 Physical manipulation may have advantages over indirect interaction

There does seem to be a benefit to the physical manipulation of objects on a tabletop space (i.e., PUI or TUI) over mouse driven interaction in a 2D space for a spatial task. Statistical analysis reported subject's higher level of engagement on a tabletop space (i.e., PUI or TUI) than on a non-tabletop space (i.e., GUI). Through observation I found that some children had difficulty rotating GUI based puzzle pieces. Observational analysis revealed that children were much more active with body movement in both the physical (traditional) and tangible conditions. I also observed evidence of moving the body to engage in perspective taking in both physical (traditional) and tangible condition, but of course not for the GUIs.

6.7.5 Interface style affects collaboration style

I observed different collaboration styles used by children on different interface styles. Children use an independent parallel play strategy on physical (traditional) and tangible user interfaces, but use a sequential turn-taking strategy on graphical user interfaces. The multiple access points afforded by a tabletop game (tangible and traditional) combined with enough space to move supported parallel independent play rather than sequential turn taking. This result is also in line with Scott's (2003) suggestion that users interact concurrently when the collaborative medium supports it.

6.7.6 Gender composition of pairs affects children's enjoyment and engagement depending on interface style

Fails (2005) indicated that gender may affect the type of interaction with the different interface environments. Statistical analysis results of this study indicated pairs of children's enjoyment and engagement were affected differently by different gender pairings across the three interfaces. I suggest that interface designers should consider gender disparities when designing tangibles for children.

6.8 Summary

One aim of this experiment was to generate a set of design guidelines that could be implemented in the design of tangible user interfaces in the future. The research facilitates my observations of many interesting trends. However, the results still must be interpreted with caution. Analysis results of enjoyment indicated that users reported similar enjoyment levels over the three interface styles. Analysis results of engagement revealed difference of user's subsequent play time, numbers of repeat play times and completion times existed between PUI and GUI, and between TUI and GUI. However those factors were similar on PUI and TUI. The result suggests that the benefit of physical manipulation could contribute to user's engagement on puzzle solving tasks. The result of collaboration analysis suggested several implications for designing tangibles for children. There is solid evidence that the gender composition of pairs has a significant effect on children's enjoyment depending on interface styles. This information provides valuable insights into the design of tangible interaction for children.

CHAPTER 7: CONCLUSION

7.1 Overview

The comparative design of this study allowed me to elicit and reflect on the fundamental differences on children's enjoyment, engagement and collaboration on three different interface styles. I found evidence to support the benefit of tangible tabletop designs for collaborative problem solving activities. I found that children's self-reports of enjoyment were similar for all three interface styles. I also found that children took longer and had more difficulty completing puzzles in the GUI condition. The finding that significantly more pairs in the PUI and TUI conditions engaged in repeat play, starting the puzzle a second time, is proposed as a better indication of engagement than time to the first completion. I also found significant gender pairing effects on subject's enjoyment and engagement across three interface styles.

In summary, this study contributes knowledge to the growing number of empirical studies comparing attributes of tangible, graphical and physical (traditional) user interfaces for school age children. The findings of this study are a springboard for continued comparative research on physical, graphical and tangible user interfaces for playful learning tasks for young children.

In this chapter, based on the findings and discussion presented in the previous chapters, I discussed several limitations of the design of this study (Section 7.2), and propose the direction and possibility for the future study (Section 7.3).

7.2 Limitations

There are several limitations of the study. First, a more controlled comparison would involve a comparable display size for the GUI condition. In this study, I used a traditional laptop to deliver the GUI puzzles. The size of the display screen restricted the size of the GUI puzzle pieces per se, which could interfere with the comparability across the three implementations.

Second, the operationalization of engagement for children remains problematic. Task time may more accurately measure cognitive difficulty than engagement in problem solving games. Time-related measurement is one kind of approach to evaluate engagement. Therefore, alternative methods suggested by previous relevant research, such as coding video footage according to a set of positive and negative instantiations (Hanna, et al., 1997; Read et al., 2000), could be used for triangulating the analysis on engagement, although the measures are very subjective. Multiple measures are definitely needed.

Third, I suggest that dual-cursor technology should be included for control purposes, although doing so might reduce ecological validity since dual mouse computer applications are rare. Through study observation, although I didn't observe a significant conflict in children's collaborative interaction in the GUI condition, the sequential-turn-taking strategy still might cost extra time in communication, including role switching and waiting for the input device. Scott et al. (2003) proposed that when concurrent, multi-user interaction is supported on a shared display, children exhibit collaborative behaviour similar to their interactions during paper-based activities. Thus, I suggest that dual-cursor technology could be developed for future explorations.

Furthermore, future studies should involve other forms of games. So far, the design of this study has been limited to one type of game only, the puzzle, and I suggest that some other playful learning tasks or spatial games, such as whack-a-mole or Tetris, could be implemented on these three interfaces, and further tested in future.

Last, the dynamical relation between enjoyment and engagement is worth exploring further. Enjoyment and engagement as the two pivotal dependent variables were analyzed systematically in this research study. Both of them were examined under several interrelated measurements. However, this study limited the scope of measuring these two variables separately. The interrelation between enjoyment and engagement was not examined. This suggests the future study continuing to explore whether and how enjoyment is related to engagement, which is in line with the viewpoints of several psychological researchers (e.g., Csikszentmihalyi, 1990; Markopoulos, 2003).

7.3 Future Research

There are many permutations of measures, tasks, interfaces styles, environments and user grouping that could experimentally tested. The experimental design limitations (addressed previously) could be immediately addressed by fine-tuning the design of this study.

For future research, I suggest that this study could be tested with older children (approximate 9 to 14 years old) or even adults. During the experimental study at Science World, I observed that a large number of older children (aged from 10 to 14 years old) showed high enthusiasm to participate in the TUI and GUI conditions. There were many adult users, even, who demonstrated their strong interest to solve the TUI puzzles during their stays.

I also suggest using an open ended task as a partial solution to how to measure engagement using time. Instead of giving pairs of children 15 minutes allocated time, I would let users decide when to finish their tasks by themselves. I would compare how long children play on an open ended task in each condition.

In summary, this study takes an important step forward by comparing three different user interfaces for evaluating school-aged children's enjoyment, engagement and collaboration on a playful learning task. This study is known as the first comparative study, which compared physical (traditional), graphical and tangible user interfaces for school-aged children in this research field. It contributes significantly to the existing body of designing tangible interaction for children. Instead of following the current trends by building new tangibles for young users, the design of this study raises the question of why we need to design tangibles for children. The second contribution of this study is the development of an extensible tabletop prototype, which uses fiducial markers and a camera vision system to track user driven events. The third contribution is the set of design considerations for the development of enjoyable and engaging tangibles. However, there are still many aspects and factors that could be discussed and further explored. I hope this study can guide the development of tangible interface systems for children in an optimal way. Hence, tangible developers can create more practical, valuable and innovative tangible systems for children in the future.

APPENDICES

Appendix A User Study Timelines and Scripts

PUI Session: Total (estimated) **time**: 30 minutes

Session 1 – PUI based jigsaw puzzle game			
Time	**Action**	**Description**	**Data Type**
4:00	Intro & pre-questionnaire	Start at regular table with pencils & sheets. All sheet headers should be filled in. Welcome: Introduce yourself and ask children their names. You will be playing a jigsaw puzzle game today. Before we start, I will ask you to answer some questions which will help us know a little about you. Here is the answer sheet. I will read each question out loud for you. Please put an X or circle one smiley face answer that is best for you. There are five kinds of smiley faces you can use to tell us how true something is for you. The saddest is "not true at all" then "not very true" then "somewhat true" then "true" then the most smiley is "very true". So if we said "I like ice cream" you might circle the true or very true smiley face (if you do like ice cream). Read pre-Questionnaire questions to them.	Pre-Questionnaire
1:00	Intro	Start at table with laptop. Ok, Great! Now we'll do a puzzle. Have you ever played any puzzles on a computer before? You're going to play a cardboard jigsaw puzzle on this table together.	None

		You two will have 15 minutes to play with the puzzle. You can do it as many times as you like. When you're done playing with it you can wait over with the books. When the total 15 minutes is done, we will let you know. Then we will ask you some more smiley face questions about your feelings and experience with the puzzle. Before we start, I will show you how to play this game. Demonstrate 1. The reference poster (Here we prepared a poster for you) 2. You can decided whether to use the poster as a reference (put on the floor, on the bench or anywhere you like) or to build your puzzle on top of this poster. Do you have any questions? Go ahead and start.	
15:00	Task 2 Boat	Start video camera. Start timer on watch. Leave children play the puzzle by themselves. Record observational notes and number of times they start and stop puzzle on sheet.	Notes sheet Watch/Timer Video
1:00	Close up	When children done turn timer off and record elapsed time. **If children say they're done Prompt:** If you want you can do it again or you can read some books until the 15 minutes is up. Timer off at 15 minutes. Ok, time is up. It's time to ask you questions. Close laptop. Video off. Video off. Move video camera to tangible tabletop. **Prompt (Puzzle):** Ok, finish up now. It's time to stop. You can finish this puzzle later, but you need to complete all these smileyface questions first. **Prompt (Books):** Ok, it's time to stop. Please come with me. (repeat)	Make sure you record time, times started, times completed and take observational notes as per protocol. For those "at books", when time is up, ask them come back for the questions.
6:00	Post-questionnaire	At table Here is your answer sheet. As we did before, I will read the question to you. For each of the statements, you will use the smiley face to indicate	Post-Questionnaire + Audio Recorder

		how true it is for you. Please circle just one of the faces for each question. Do Smiley questions. Next, do open ended questions. Switch Audio on. Add header to audio with date, session, interface, castle (i.e., 1st part of session). Now I will ask you two more questions and you can each just tell me your answer out loud. Make sure you give both a chance to answer. **Prompt:** Tell me more about that. **Prompt:** Is there anything else?	
8:00	Finish & Clean up stations	Great! You have finished the first part. Good job! Thank you very much for your cooperation. We really appreciate your help for our study.	

GUI Session: Total (estimated) time: 30 minutes

Session 2 – GUI based jigsaw puzzle game			
Time	**Action**	**Description**	**Data Type**
4:00	Intro & pre-questionnaire	Start at regular table with pencils & sheets. All sheet headers should be filled in. Welcome: Introduce yourself and ask children their names. You will be playing a jigsaw puzzle game today. Before we start, I will ask you to answer some questions which will help us know a little about you. Here is the answer sheet. I will read each question out loud for you. Please put an X or circle one smiley face answer that is best for you. There are five kinds of smiley faces you can use to tell us how true something is for you. The saddest is "not true at all" then "not very true" then "somewhat true" then "true" then the most smiley is "very true". So if we said "I like ice cream" you might circle the true or very true smiley face (if you do like ice cream). Read pre-Questionnaire questions to them.	Pre-Questionnaire

1:00	Intro	Start at table with laptop. Ok, Great! Now we'll do a puzzle. Have you ever played any puzzles on a computer before? You're going to play a jigsaw puzzle on this computer. Since we only have one mouse, you will have to share when you play. You two will have 15 minutes to play with the puzzle. You can do it as many times as you like. When you're done playing with it you can wait over with the books. When the total 15 minutes is done, we will let you know. Then we will ask you some more smiley face questions about your feelings and experience with the puzzle. Before we start, I will show you how to play this game. Demonstrate 1. How to move a piece (You can move a piece like this) 2. How to connect pieces (You can connect like this) 3. How to rotate (You turn pieces like this) 3. How to show/hide the source image (You can turn off the image like this) 4. How to restart the game (You can start to game over like this. Do you have any questions? Go ahead and start.	None
15:00	Task 2 Boat	Start video camera. Start timer on watch. Leave children play the puzzle by themselves. Record observational notes and number of times they start and stop puzzle on sheet.	Notes sheet Watch/Timer Video
1:00	Close up	When children done turn timer off and record elapsed time. **If children say they're done Prompt:** If you want you can do it again or you can read some books until the 15 minutes is up. Timer off at 15 minutes. Ok, time is up. It's time to ask you questions.	Make sure you record time, times started, times completed and take observational notes as per protocol. For those "at books", when time is up, ask them come back for the questions.

		Close laptop. Video off. Video off. Move video camera to tangible tabletop. **Prompt (Puzzle):** Ok, finish up now. It's time to stop. You can finish this puzzle later, but you need to complete all these smileyface questions first. **Prompt (Books):** Ok, it's time to stop. Please come with me. (repeat)	
6:00	Post-questionnaire	At table Here is your answer sheet. As we did before, I will read the question to you. For each of the statements, you will use the smiley face to indicate how true it is for you. Please circle just one of the faces for each question. Do Smiley questions. Next, do open ended questions. Switch Audio on. Add header to audio with date, session, interface, castle (i.e., 1st part of session). Now I will ask you two more questions and you can each just tell me your answer out loud. Make sure you give both a chance to answer. **Prompt:** Tell me more about that. **Prompt:** Is there anything else?	Post-Questionnaire + Audio Recorder
8:00	Finish & Clean up stations	Great! You have finished the first part. Good job! Thank you very much for your cooperation. We really appreciate your help for our study.	

TUI Session: Total (estimated) **time**: 30 minutes

Session 3 – TUI based jigsaw puzzle game			
Time	**Action**	**Description**	**Data Type**
4:00	Intro & pre-questionnaire	Start at regular table with pencils & sheets. All sheet headers should be filled in. Welcome: Introduce yourself and ask children their names. You will be playing a jigsaw puzzle game today. Before we start, I will ask you to answer some questions which will help us know a little about you.	Pre-questionnaire

		Here is the answer sheet. I will read each question out loud for you. Please put an X or circle one smiley face answer that is best for you. There are five kinds of smiley faces you can use to tell us how true something is for you. The saddest is "not true at all" then "not very true" then "somewhat true" then "true" then the most smiley is "very true". So if we said "I like ice cream" you might circle the true or very true smiley face (if you do like ice cream). Read pre-Questionnaire questions to them.	
1:00	Intro & Demo	Ok, Great! Now we'll do a puzzle. Move children to TUI tabletop. Activate table with mouse so image shows. Reset with reset block if necessary. At this station, you can see the table. This table has a computer in it. The puzzle image is shown on the table top. During your play, whenever you get the right pieces connected together, the table will make a "laser" sound. The computer may give you some other kinds of feedback that let you know you're making progress, but you have to figure that out by yourself. **Show reset block.** This is a reset block. If you want to start over at any time just clear off the pieces and put this blocks with the pattern down on the table. **Show picture on/off block.** This block turns the image off. You can put it pattern side down to turn the picture off at any time you like. You will do the jigsaw puzzle on the table together. You will have 15 minutes to play with the puzzle. You can do it as many times as you like. When you're done playing with it you can wait over with the books.	None

		When the total 15 minutes is done, we will let you know. Then we will ask you some more smiley face questions about your feelings and experience with the puzzle. Ok! Before we start, do you have any questions?	
15:00	Task 1 Castle	Start video camera. Start timer on watch. Leave children play the puzzle by themselves. Record observational notes and number of times they start and stop puzzle on sheet.	Record notes on Notes sheet Timer watch – record on sheet. Video.
1:00	Close up	When children are done, turn off timer and record elapsed time. **If children say they're done Prompt:** If you want you can do it again or you can read some books until the 15 minutes is up. Timer off at 15 minutes. Ok, time is up. It's time to ask you questions. We'll go back to the table. Video off. Move video camera to table with laptop. Move back children to table with laptop. **Prompt (Puzzle):** Ok, finish up now. It's time to stop. Please come with me. (repeat) **Prompt (Books):** Ok, it's time to stop. Please come with me. (repeat)	For those "at books", when time is up, ask them come back for the questions
6:00	Post-questionnaire	At Laptop table Here is your answer sheet. As we did before, I will read the question to you. For each of the statements, you will use the smiley face to indicate how true it is for you. Please circle just one of the faces for each question. Do Smiley questions. Next, do open ended questions. Switch Audio on. Add header to audio with date, session, interface, castle (i.e., 1st part of session). Now I will ask you two more questions and you can each just tell me your answer out loud. Make sure you give both a chance to answer.	Post-Questionnaire + Audio Recorder

		Prompt: Tell me more about that. **Prompt:** Is there anything else?	
8:00	Finish & Clean up stations	Great! You have finished the first part. Good job! Thank you very much for your cooperation. We really appreciate your help for our study.	

Appendix B Pre-questionnaire

Instructions: For each question, choose the best answer for you by putting an X in the box or circling the face.

1. How old are you?

   ☐ 5-6 ☐ 7-8 ☐ 9-10 ☐ older than 10

2. I am a ☐ boy ☐ girl

3. Are you a fluent English speaker?

   ☐ YES. I am ☐ NO. I'm not. ☐ Somewhat.

4. How often do you use the computer(s) at your home or your school?

   ☐ never ☐ about once a month ☐ about once a week

   ☐ about once a day or more

5. I feel comfortable to use a computer mouse.

Not at all true | Not very true | Somewhat true | True | Very true

6. Have you ever done a jigsaw puzzle before?

   ☐ Yes ☐ No

Instructions: For each sentence, choose how true it is for you by circling one face.

7. I like doing jigsaw puzzles.

Not at all true | Not very true | Somewhat true | True | Very true

8. I'm interested in wizard, witches and magic.

Not at all true | Not very true | Somewhat true | True | Very true

9. I'm interested in pirates, treasure and adventures.

Not at all true | Not very true | Somewhat true | True | Very true

Appendix C Post-questionnaire

Instructions: For each question, choose the best answer for you by circling the face.

1. While I was working on this puzzle I was thinking about how much I enjoyed it. (Interest/enjoyment)

2. I did not feel at all nervous about doing this puzzle. (Pressure/tension) (R)

   Nervous: It is a feel like when you scared about a test, or standing in front of class to talk. You feel sick in your stomach

3. I felt that it was my choice to do this jigsaw puzzle. (Perceived choice)

4. I think I am pretty good at solving this puzzle. (Perceived competence)

5. I found doing this puzzle very interesting. (Interest/enjoyment)

6. I felt tense while doing the puzzle. (Pressure/tension)

   Tense: Your whole body feel like this way. Something that you feel tight (crouch your body)

7. I think I did pretty well at this puzzle, compared to others. (Perceived competence)

8. Doing this puzzle was fun. (Interest/enjoyment)

9. I felt relaxed while doing this puzzle. (Pressure/tension) (R)

   Relaxed: you feeling calm in your body, like the way you fall sleep. It's opposite of feeling worry.

10. I enjoyed doing this puzzle very much. (Interest/enjoyment)

11. I didn't really have a choice about doing the puzzle. (Perceived choice) (R)

12. I am satisfied with my performance at this puzzle. (Perceived competence)

13. I was anxious while solving this puzzle. (Pressure/tension)

    Anxious: means worried

14. I thought the jigsaw puzzle was very boring. (Interest/enjoyment) (R)

15. I felt like I was doing what I wanted to do while I was working on the puzzle. (Perceived choice)

16. I felt pretty skilled at jigsaw puzzles. (Perceived competence)

    Skilled: means good at that

17. I thought doing this puzzle was very interesting. (Interest/enjoyment)

18. I felt pressured while doing this puzzle. (Pressure/tension)

    Pressured: a feel like when you have to do something quickly but you don't have time

19. I felt like I had to do this puzzle. (Perceived choice) (R)

20. I would describe this puzzle as very enjoyable. (Interest/enjoyment)

21. I did this puzzle because I had no choice. (Perceived choice) (R)

22. After working at this puzzle for awhile, I felt pretty competent. (Perceived competence)

    Competent means it is easy for you because you are good at it.

23. I would like to do another puzzle like this one with one of my good friends.

24. I like the way I can move the puzzle pieces around while I solve the puzzle.

25. I like the way the picture helps me do the puzzle.

26. I would like to do a similar puzzle like this by myself next time.

27. I like knowing when I've connected the pieces correctly.

28. I would like to do another puzzle like this one again some time.

    ** Before asking the following two questions, please make sure you have **turned on** the audio recorder.

    Give the following information at the beginning of each session, including test date, session number, first or second part of the session, name of the facilitator, and the interface type.

29. Tell me what things you most liked about doing this jigsaw puzzle.

30. Tell me what things you most didn't like about doing this jigsaw puzzle.

REFERENCE LIST

Africano, D., Berg, S., Lindbergh, K., Lundholm, P., Nilbrink, F., and Persson, A. Designing tangible interfaces for children's collaboration. In *CHI 2004 Extended Abstracts on Human Factors in Computing Systems*, ACM Press (2004), 853-868.

Aliakseyeu D., Martens J.B., Subramanian S., Vroubel, M. and Wesselink, W. Visual interaction platform. In *Proceedings of the Human-Computer Interaction INTERACT 2001*, IOS Press (2001), 232-239.

Andersen, K. 'Ensemble': playing with sensors and sound. In *CHI 2004 Extended Abstracts on Human Factors in Computing Systems*, ACM Press (2004), 1239-1242.

Antle, A. N. The CTI framework: informing the design of tangible systems for children. In *Proceedings of the 1st international Conference on Tangible and Embedded interaction*, ACM Press (2007), 195-202.

Bakeman, R., & Brown, J.V. Early interaction: consequences for social and mental development at three years. Child development, Vol. 51, (1980), 437-447.

Berman, R. Preschool knowledge of language: what five-year olds know about language structure and language use. In *C. Pontecorvo (ed.) Writing Development: an Interdisciplinary View*, Amsterdam: John Benjamin Publishing, (1977), 61-67.

Bohn, Jürgen. The Smart Jigsaw Puzzle Assistant: Using RFID Technology for Building Augmented Real-World Games. Workshop on *Gaming Applications in Pervasive Computing Environments at Pervasive 2004*, (2004).

Brennan, D. and Bleakley, E.W. 'Predictors, Problems and Policies for Post School Participation'. In *J. Kremer, K. Trew and S. Ogle (eds) Young People's Involvement in Sport*, London: Routledge (1997), 78–97.

Brophy, J.E., & Good, T.L. *Teacher-student relationships: causes and consequences.* New York, Holt, Rinehart and Winston, (1974).

Brustad, R. Who Will Go Out and Play? Parental and Psychological Influences on Children's Attraction to Physical Activity. In *Paediatric Exercise Science5*, (1993), pp. 210–223.

Bruner, J., Going beyond the information given. In *H. Gruber, G. Terrell, & M. Wertheimer (Eds.), Contemporary approaches to cognition*, Cambridge, MA: Harvard University Press (1973), 258-290.

Carroll, B. and Loumidis, J. Children's Perceived Competence and Enjoyment in Physical Education and Physical Activity Outside School. In *European Physical Education Review 7(1)* (2001), 24–43.

Cassell, J. and Jenkins, H. Chess for girls? Feminism and computer games. In *From Barbie To Mortal Kombat: Gender and Computer Games*, *J. Cassell and H. Jenkins, Eds*. MIT Press (1998), Cambridge, MA, 2-45.

Clements, D., Playing with computers, playing with ideas. In *Educational Psychology Review 7 (2)*, (1995), 203–207.

Crook, C. Children as computer users: the case of collaborative learning. In *Computers Education*, vol. 30, No.3/4, Great Britain, (1998), 237-247.

Csikszentmihalyi, M., *Flow: The Psychology of the Optimal Experience*. New York: HarperCollins, (1990).

Davis, W. A causal theory of enjoyment. In *Proceedings of Mind 91, XCI Press* (1982), 240-256.

Deci, E.L. and Ryan, R.M. Intrinsic Motivation and Self-Determination in Human Behaviour. New York Plenum Press (1985).

Deci, E. L., Ryan, R. M., Self-Determination Theory and the Facilitation of Intrinsic Motivation. In *Social development, and Well-being. American Psychologist 55, 1* (2000), 68-78.

Deguchi, A., Sugimoto, M., Yamamoto, T., Yamaguchi, E., Kusunoki, F., Seki, T., Inagaki, S., Tachibana, S., and Takeuchi, Y. CarettaKids: a system for supporting children's face-to-face collaborative learning by integrating personal and shared spaces. In *Proceedings of the 2006 Conference on interaction Design and Children*, *IDC '06,* ACM Press (2006), New York, NY, 45-48.

Dix, A. 2003. Being playful: learning from children. In *Proceeding of the 2003 Conference on interaction Design and Children*, *IDC '03.,* ACM Press (2003), New York, NY, 3-9.

Doise, W. and Mugny, G. *The social Development of the Intellect.* Pergamon Press, Oxford. (1984).

Donker, A. and P. Markopoulos. A comparison of Think-aloud, Questionnaires and Interviews for testing usability with children. In *The 16th BCS British-HCI 2002* (2002), London, UK, 305-316.

Druin, A. *The Design of Children's Technology*. Morgan Kaufmann Publishers Inc. (1998).

Druin, A. The role of children in the design of new technology. In *Behaviour & Information Technology, vol. 21, No. 1*, (2002), 1-25.

Edwards, D., & Mercer, N. Common knowledge: The development of understanding in the classroom. London, Methuen, Rutledge, (1987).

Fails, J. A., Druin, A., Guha, M. L., Chipman, G., Simms, S., and Churaman, W. Child's play: A comparison of desktop and physical interactive environments. In *Proceeding of the 2005 Conference on Interaction Design and Children*, ACM Press (2005), 48-55.

Fernaeus, Y. and Tholander, J. Finding design qualities in a tangible programming space. In Proceedings of the SIGCHI Conference on Human Factors in Computing System,s. R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. CHI 2006, ACM Press (2006), 447-456.

Fishkin, K. P., Moran, T. P., and Harrison, B. L. Embodied User Interfaces: Towards Invisible User Interfaces. In *Proceedings of EHCI'98*, (1998), Heraklion, Greece.

Fishkin, K.A. Taxonomy for and Analysis of Tangible Interfaces. In *Personal and Ubiquitous Computing 8 (5)*, (2004), 347-358.

Fitzmaurice, G.W., Ishii, H. & Buxton, W. Bricks: Laying the Foundations for Graspable User Interfaces. In Proceedings of CHI'95, ACM Press (1995), 442-449.

Forman, E. The role of peer interaction in the social construction of mathematical knowledge. In *International Journal of Educational Research*, Vol. 13, (1989), 55-69.

Gardner, H. J., and Martin, M. A. Analyzing Ordinal Scales in Studies of Virtual Environments: Likert or Lump it. In *Presence. Vol. 16. No. 4,* (2007), 439-446.

Goffman, J.M., & Bakeman, R. The sequential analysis of observational data. In M.E. Lamb, S.J. Suomi, & C.R. Stephenson (Eds.), *Social interaction analysis: methodological issues*. Madison: University of Wisconsin Press, (1979)

Hanna, L., Risden, K. and Alexander, K. Guidelines for Usability Testing with Children. In *Interactions, 1997. 5*, (1997), 9-14.

Hanna, L., Risden, K., Czerwinski, M., & Alexander, K. J. The role of usability research in designing children's computer products. In *The design of children's technology 1998*, Morgan Kaufmann Publishers Inc (1998), 3-26.

Hanna, L., Neapolitan, D., and Risden, K. Evaluating computer game concepts with children. In *Proceeding of the 2004 Conference on interaction Design and Children: Building A Community*, ACM Press IDC (2004), 49-56.

Hatano, G., & Inagaki, K. Desituating cognition through the construction of conceptual knowledge. In P. Light & G. Butterworth (Eds.), *Context and cognition: ways of learning and knowing*. London, Harvester-Wheatsheaf. (1992), 115-133.

Healy, J. M. *Failure to Connect: How Computers Affect Our Children's Minds*. Simon and Schuster, New York, NY, USA, (1998).

Heidegger, M. *Being and Time* (J. Macquarrie & E. Robinson, Trans.). Oxford: Blackwell, (1990).

Hinckley, K., Pausch, R., Goble, J., and Kassel, N. Passive Real-World Interface Props for Neurosurgical Visualization. In *Proceedings of CHI '94*, (1994),452-458.

Holmquist, L., Schmidt, A. and Ullmer, B. Tangible interfaces in perspective: Guest editors' introduction. In *Personal and Ubiquitous Computing 8, 5* (2004), 291-293.

Hornbæk, K. Current Practice in Measuring Usability: Challenges to Usability Studies and Research. *Int. Journal of Human-Computer Studies 64, 2* (2006), 79-102.

Hornecker, E. and Buur, J. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of CHI 2006*, ACM Press (2006), 437-446.

Hoyles, C. What is the point of group discussion in mathematics Studies in Mathematics, 16, (1985), 205-214.

Hughes, J. Assessment of children's social competence. In Reynolds, C.R. and Kamphaus, R. Editors, Handbook of psychological and educational assessment of children, (1990), Guilford, New York, 423-444.

Hymel, S., Zinck, B., & Ditner, E. Cooperation versus competition in the classroom. *Exceptionality Education Canada*, 3:1-2. (1993), 103-128.

Inkpen, K., Booth, K. S., Klawe, M., and Upitis, R. *Cooperative Learning in the Classroom: the Importance of a Collaborative Environment for Computer-Based Education*. Technical Report. UMI Order Number: TR-94-05., University of British Columbia. (1994)

Inkpen, K., Booth, K.S., Klawe, M. and Upitis, R. Playing together beats playing apart, especially for girls. In *Proceedings of Computer-Supported Collaborative Learning (CSCL '95)*, (eds. J. Schnase & E. Cunnius), Lawrence Erlbaum, Hillsdale, NJ. (1995), 252-259.

Inkpen, K., McGrenere, J., Booth, K. S., and Klawe, M. The effect of turn-taking protocols on children's learning in mouse-driven collaborative environments. In *Proceedings of the Conference on Graphics interface '97* (Kelowna, British Columbia, Canada). W. A. Davis, M. Mantei, and R. V. Klassen, Eds. Canadian Information Processing Society, Toronto, Ont., Canada, (1997), 138-145.

Inkpen, K. M., Ho-Ching, W., Kuederle, O., Scott, S. D., and Shoemaker, G. B. This is fun! We're all best friends and we're all playing: supporting children's synchronous collaboration. In *Proceedings of the 1999 Conference on Computer Support For Collaborative Learning*, ACM Press (1999), 252-259.

Ishii, H., Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In *Proceedings of CHI 97*, ACM Press (1997), 234-241.

Jacques, R., Preece, J., Carey, T. Engagement as a Design Concept for Multimedia. In *Canadian Journal of Educational Communications 24, 1*, (1995), 49-59.

Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international Conference on Tangible and Embedded interaction TEI 2007*, ACM Press (2007), 139-146.

Klemmer, S. R., Hartmann, B., and Takayama, L. How bodies matter: five themes for interaction design. In *Proceedings of the 6th Conference on Designing interactive Systems (University Park, PA, USA, June 26 - 28, 2006). DIS '06.* ACM Press (2006), New York, NY, 140-149.

Khandelwal. M. *Teaching table: a tangible mentor for pre-kindergarten math education*. School of literature, communication and culture, Ivan Allen College, Georgia Institution of Technology. (2006)

Kremer, J., Trew, K. and Ogle, S. *Young People's Involvement in Sport*. London: Routledge, (1997).

MacFarlane, S., Read, J., Höysniemi, J., and Markopoulos, P. Tutorial: Evaluating interactive products with and for children. In *Conference of Human Factors in Computing Systems , CHI 2004*. (2004), Vienna, Austria, 1-5.

Malone, T.W. What makes things fun to learn? Heuristics for designing instructional computer games. In *Proceedings of the 3rd ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems*. Palo Alto, California, United States, SIGSMALL '80. ACM, New York, NY, (1980), 162-169.

Malone, T.W. Toward a Theory of Intrinsically Motivating Instruction. In *Cognitive Science 4*, (1981), 333-369.

Malone, T.W., & Lepper, M.R. Making learning fun: A taxonomy of intrinsic motivations for learning. In *R. E. Snow & M. J. Farr (Eds.), Aptitude, Learning and instruction*. Hillsdale (1987), NJ, Erlbaum.

Mandigo, J. L. & Holt, N. L. Putting theory into practice: how cognitive evaluation theory can help us motivate children in physical activity environments. In *Journal of Physical Education, Recreation, and Dance, 71(1)*, (2000), 44-49.

Markopoulos, P., Bekker, M. Interaction Design and Children. In *Interacting with Computers 15*, (2003), 141-149.

Marshall, P. Do tangible interfaces enhance learning? In Proceedings of the 1st international Conference on Tangible and Embedded interaction TEI 2007, ACM Press (2007), 163-170.

Mazalek, A. and Davenport, G. A tangible platform for documenting experiences and sharing multimedia stories. In *Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence (Berkeley, California). ETP '03,* ACM Press (2003), New York, NY, 105-109.

McNerney, T. S. From turtles to Tangible Programming Bricks: Explorations in physical language design. In *Personal Ubiquitous Computing 8, 5*, (2004), 326-337.

Mercer, N. The quality of talk in children's collaborative activity in the classroom. In *Learning and instruction*, Vol. 6, No. 4, Elsevier Science Ltd., (1996), 359-377.

Monk, A. Fun, Communication and Dependability: Extending the Concept of Usability. In *People and Computers XVI, Springer*, (2002), 3-14.

Monk, A., Hassenzahl, M., Blythe, M., and Reed, D. Funology: designing enjoyment. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*. Minneapolis, Minnesota, USA, CHI '02. ACM, New York, NY. (2002), 924-925.

Montemayor, J., Druin, A., Farber, A., Simms, S., Churaman, W., and D'Amour, A. Physical Programming: Designing Tools For Children To Create Physical Interactive Environments. In *Proceedings of CHI 2002*, ACM Press (2002), 299-305.

Montessori, M. Montessori Spontaneous Activity In Education: The Advanced Montessori Method. John Wiley, New York, (1965).

Papert, S. Mindstorms: Children, computers, and powerful ideas. New York, NY, USA, (1980).

Parten, M.B. Social participation among preschool children. In *Journal of Abnormal and Social Psychology, 27*. (1932), 243-269.

Piaget, J. *The Moral Judgement of the Child*. Routledge and Kegan Paul, London, (1932).

Piaget, J. Play, dreams and imitation in childhood. New York: Norton. (1962)

Prensky, M. Fun, Play and Games: What Makes Games Engaging. In *Digital game-based learning*. Comput. Entertain. (2001).

Price, S., Rogers, Y., Scaife, M., Stanton, D. and Neale, H. Using 'tangibles' to promote novel forms of playful learning. In *Proceedings of the Interacting with Computers 15, 2*, ACM Press (2003), 169-185.

Price S., Rogers Y. Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. In *Computers and Education, 43,* (2004), 137-151.

Raffle, H., Parkes, A., Ishii, H., and Lifton, J. Beyond record and play: Backpacks: tangible modulators for kinetic behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. CHI 2006*, ACM Press (2006), 681-690.

Raffle, H., Ishii, H., and Yip, L. Remix and Robo: sampling, sequencing and real-time control of a tangible robotic construction system. In *Proceedings of the 6th international Conference on interaction Design and Children*, *IDC '07,* ACM Press (2007), New York, NY, 89-96.

Rauterberg, M., Fjeld, M., Krueger, H., Bichsel, M., Leonhardt, U. & Meier, M. BUILD-IT: A video based interaction technique of a planning tool for construction and design. In *Video program of ACM CHI 1998,* ACM Press (1998), 177-178.

Read, J. C., MacFarlane, S. J. and Casey, C. Endurability, engagement and expectations: Measuring children's fun. In *Proceedings of the Interaction Design and Children*, Shaker Publishing (2002), Germany, 189-198.

Resnick, M., Martin, F., Berg, R., Borovoy, R., Colella, V., Kramer, K., and Silverman, B. Digital manipulatives: New toys to think with. In Proceedings of CHI '98, ACM Press (1998), 281-287.

Resnick, M., Bruckman, A., Martin, F. Constructional Design: Creating New Construction Kits for Kids. In *Druin, A., (Ed.), The Design of Children's Technology, Morgan Kauffman*, (1999), USA.

Risden, K., Hanna, E., & Kanerva, A. Dimensions of intrinsic motivation in children's favorite computer activities. In *Society for Research in Child Development,* (1997), Washington, DC.

Rogers, Y., Scaife, M., Gabrielli, S., Smith, H. and Harris, E.A. Conceptual framework for mixed reality environments: Designing novel learning activities for young children. In *Proceedings of Presence 11, 6,* (2002), 677-686.

Rogers, Y. and Muller, H. A framework for designing sensor-based interactions to promote exploration and reflection in play. In *Journal of Human Computer Studies 64, 1*, (2006), 1-38.

Rozendaal, M. M., Keyson, D. V., and de Ridder, H. Product behavior and appearance effects on experienced engagement during experiential and goal-directed tasks. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and interfaces DPPI '07.* ACM Press (2007), New York, NY, 181-193.

Rubin, K.H. Relation between social participation and role-taking skill in preschool children. In *Psychological Reports*, 39. (1976), 823-826.

Ryan, R.M. (2006) Intrinsic Motivation Inventory (IMI) http://www.psych.rochester.edu/SDT/measures/intrins.html

Ryan, R.M., & Deci, E. L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. In *American Psychologist, 55*, (2000) 68–78.

Salen, K., & Zimmerman, E. *Rules of play*. MIT Press (2003), Cambridge, Massachusetts.

Salomon, G. & Globerson, T. When teams do not function the way they ought to. In *Proceedings of International Journal of Educational Research 13*, (1987), 89-100.

Saracho, O.N. Young children's play behaviors and cognitive style. In *Journal of Early child development and care,* Vol. 22. (1985), 1-18.

Scanlan, T.K., Stein, G.L. and Ravizza, K. An In-Depth Study of Former Elite Figure Skaters: II. Sources of Enjoyment. In *Proceedings of Journal of Sport Exercise Psychology 11*, (1989), 65-83.

Scarlett, W. George. *Children's Play.* Thousand Oaks, CA: Sage Publications, (2004).

Schunk, D. Verbalisation and children's self regulated learning. *Contemporary Educational Psychology*, Vol.11, (1986), 347-369.

Scott S.D., Mandryk R.L., Inkpen K.M. Understanding children's collaborative interactions in shared environments, In *Proceedings of the Journal of Computer Aided Learning 19*, *2,* (2003), 220-228.

Scott S.D., Grant K.D., and Mandryk R.L. System Guidelines for Co-located, Collaborative Work on a Tabletop Display. In *Proceedings of European Conference of Computer-Supported Cooperative Work (ECSCW)*, Kluwer Academic Publishers (2003), 159-178.

Siann, G., Durndell, A., Macleod, H. and Glissov, P. Stereotyping in relation to the gender gap in computing. In *Education Research 30*, (1988), 98-103.

Sluis, R. J., Weevers, I., van Schijndel, C. H., Kolos-Mazuryk, L., Fitrianie, S., and Martens, J. B. Read-It: Five-to-seven-year-old children learn to read in a tabletop environment. In *Proceeding of the 2004 Conference on interaction Design and Children: Building A Community*, ACM Press (2004), 73-80.

Stanton, D., Bayon, V., Neale, H., Ghali, A., Benford, S., Cobb, S., Ingram, R., Wilson, J., Pridmore, T., O'Malley, C. Classroom collaboration in the design of tangible interfaces for storytelling. In *Proceedings of Human Factors in Computing Systems (CHI 2001)*, ACM Press (2001), 482-489.

Stoney, S., Oliver, R. Can higher order thinking and cognitive engagement be enhanced with multimedia? In *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, (1999).

Sugimoto, M., Hosoi, K., and Hashizume, H. *Caretta*: a system for supporting face-to-face collaboration by integrating personal and shared spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04,* ACM Press (2004), New York, NY, 41-48.

Terrenghi, L., Kirk, D., Sellen, A., and Izadi, S. Affordances for manipulation of physical versus digital media on interactive surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '07*, ACM Press (2007), New York, NY, 1157-1166.

Thomas, J. W., and Funaro, G. M. A multi-media, computer-based model for learner-directed, collaborative problem-solving. In Woolf, B., et al., eds., Working Notes of 1990 Spring Symposium Series on Knowledge-Based Environments for Learning and Teaching, Stanford University Press (1990), 68-71.

Ullmer, B. and Ishii, H. Emerging frameworks for tangible user interfaces. In *IBM Syst. J. 39, 3-4*, (2000), 915-931.

Ullmer, B, and Ishii, H. *Human-Computer Interaction in the New Millennium*, John M. Carroll, Ed., Addison-Wesley Press (2001), 579-601.

Verhaegh, J., Fontijn, W., and Hoonhout, J. TagTiles: optimal challenge in educational electronics. In *Proceedings of the 1st international Conference on Tangible and Embedded interaction*, ACM Press (2007), 187-190.

Vygotsky, L. S. Mind in Society: The Development of Higher Psychological Processes. Harvard University Press (1978). 52-91.

Wankel, L.M. The Importance of Enjoyment to Adherence and Psychological Benefits from Physical Activity. In *Proceedings of International Journal of Sport Psychology* 24, (1993), 151-69.

Weiss, M.R. Self Esteem and Achievement in Children's Sport and Physical Activity. In *D. Gould and M.R. Weiss (eds) Advances in Pediatric Sport Sciences, vol. 2, Behavioural Issues*, Champaign, IL: Human Kinetics (1987), 87-119.

Zuckerman O., Arida S., and Resnick M. Extending tangible interfaces for education: Digital Montessori-inspired manipulatives. In *Proceeding of CHI 2005*, ACM Press (2005), 859-868.