Available at www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Interacting with Computers 15 (2003) 203–225

Interacting with Computers

www.elsevier.com/locate/intcom

Using peer tutoring in evaluating the usability of a physically interactive computer game with children

Johanna Höysniemi[a,*], Perttu Hämäläinen[b], Laura Turkki[c]

[a]*Tampere Unit for Computer–Human Interaction, Department of Computer and Information Sciences, University of Tampere, Tampere FIN-33014, Finland*

[b]*Telecommunications software and multimedia laboratory, Helsinki University of Technology, P.O. Box 5400, FIN-02015 HUT Helsinki, Finland*

[c]*Software Business and Engineering Institute/Usability Group, Helsinki University of Technology, P.O. Box 5400, FIN-02015 HUT Helsinki, Finland*

Abstract

This paper presents a novel approach to usability evaluation with children called *peer tutoring*. Peer tutoring means that children teach other children to use the software that is evaluated. The basic philosophy behind this is to view software as a part of child's play, so that the teaching process is analogous to explaining the rules of a game such as hide and seek. If the software is easy to teach and learn, it is more likely that the amount of users increases in a social setting such as a school. The peer tutoring approach provides information about teachability and learnability of software and it also promotes communication in the test situation, compared to a test person communicating with an adult instructor. The approach has been applied to the development of a perceptually interactive user interface in QuiQui's Giant Bounce, a physically and vocally interactive computer game for 4–9 year old children. The results and experiences of using peer tutoring are promising and it has proved to be effective in detecting usability flaws and in improving the design of the game.
© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Children; Usability evaluation; Methodology; Peer tutoring; Perceptual user interface; Physical interaction

1. Introduction

The increasing amount of time children spend with computers and video games has raised questions about how the use of computer technology affects the lives of children.

* Corresponding author.

E-mail addresses: johanna@cs.uta.fi (J. Höysniemi); pjhamala@tml.hut.fi (P. Hämäläinen); laura.turkki@hut.fi (L. Turkki).

0953-5438/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0953-5438(03)00008-0

Some studies suggest that children's extended computer use may be linked to an increase in obesity, seizures and hand injuries, and displace other activities that have more developmental value, like organized sports or other social activities, thus having an impact on children's physical and social well-being (Subrahmanyam et al., 2000). The design of children's computer products must be careful and considerate in order to prevent computers from being physically and socially harmful to children.

Our work is part of a wider project looking at alternative ways of controlling and interacting with computers. We wish to find out what kind of possibilities physical and vocal interaction might offer to children in comparison with traditional interaction styles. We believe that when used appropriately, computers can be a positive factor in the physical and social development process of a child.

There has been little research investigating the use of perceptual control mechanisms such as computer vision or hearing that do not attach children directly to the computer. Such techniques enable children to use their whole body to interact with the computer (Bobick et al, 1999; D'Hooge and Goldsmith, 2001; Pinhanez et al., 2000; Sengupta et al., 2000). Our research is based on using inexpensive web cameras to detect the position and movement of the user's body, and microphones to detect the user's voice. This technological framework is used to develop *QuiQui*'s *Giant Bounce*, a computer game that is aimed at developing children's physical abilities.

When designing a physically interactive and collaboration enabling computer game for children, usability studies must be carried out to ensure that the game fulfils the requirements set for physical appropriateness and social interaction. In this paper, we introduce a new method used in children's usability research, *peer tutoring approach*, that provides us with information about how children play the game physically and vocally in an authentic social context with their age mates. Simply put, the peer tutoring approach for usability evaluation means that children teach other children how to use the software that is being evaluated. Our method is based on a well known approach in children's education, and further developed in the process of designing QuiQui's Giant Bounce. In addition to peer interaction studies (Rogoff, 1990; Topping, 1988; Goodblad and Hirst, 1989) our approach has benefited a lot from traditional usability research (Nielsen, 1993; Dumas and Redish, 1993), and especially the studies and guidelines (Hanna et al., 1999, 1997; Druin et al., 1999; Cassell and Ryokai, 2001) made for the design of children's computer products.

This paper first introduces the main features of the QuiQui's Giant Bounce computer game and what is required of the usability evaluation methodology. Then, the key characteristics of the peer tutoring approach are discussed and how the approach is applied in our research. In Section 6, our observations and results of using peer tutoring are presented.

2. QuiQui's giant bounce

The peer tutoring method was developed for the usability evaluation purposes of the physically and vocally interactive computer game called QuiQui's Giant Bounce.

QuiQui's Giant Bounce is a computer game for 4–9 year old children that is controlled through movement and voice. The user interface is designed around a series of non-violent game tasks that each have a different theme of moving, for example, jumping, flying or swimming. The children are not physically attached to the computer which allows free movement in the playspace. The child controls a cartoon-like animated dragon, QuiQui, by moving in the way described in the story. For example, upon arrival to the desert playing field, the player needs to fly the QuiQui character in the sky by waving his or her hands and flexing the body. A little scream makes QuiQui let out a fiery breath to drive away distractions in the game.

2.1. Contextual movements and imitation of an avatar

The story context is an important factor in designing intuitive movements. For example, when QuiQui needs to fly through rain clouds, most of the children start to flap their hands as if they were wings. Using these *contextual movements* decreases the need for instructions and makes the learning phase shorter. Learning the movements easily is important, since children want to be in control of their world as quickly as possible (Druin, 1999; p. 67). It also keeps them interested in the game longer. Each game task has a set of contextual movements that are selected according to the milestones of children's physical development, and through usability research.

In addition to using contextual movements, the game controls are based on *avatar imitation*. In our game this means that QuiQui tries to imitate a child's movements or use of voice continuously as shown in Fig. 1. The avatar animations give hints to the children on what kind of movements are expected. The design of the imitation consists of two parts: (1) the way children find it natural and physically appropriate to control the avatar and (2) how the avatar responds to these movements within the limits of computer vision or hearing technology.

2.2. Two layers of the user interface

The user interface of QuiQui's Giant Bounce consists of two layers: (1) the human–avatar and (2) the avatar–virtual world layers. A visualization of the real world appears at

Fig. 1. An example of avatar imitation—QuiQui exhales sparkles when Eemeli shouts.

the bottom left corner of the screen (Fig. 1), where the user can see a small display of the video the webcam sends to the computer. Using this display as a guide, adults and older children can adjust the position of the camera so that children can see themselves during the game play. The interaction between the avatar and the virtual world is event-based and takes place, for example, when the avatar flies through a blue cloud and makes the cloud rain. The child controls the movements of the avatar and tries to get the avatar to interact with the objects in the virtual world. In the HCI field the research focus has been mainly on the study of interaction in the virtual domain such as information design and cognitive workload. However, our research is currently focused more on designing and evaluating the actual physical and vocal interface between the user and the avatar.

2.3. Collaborative technology

The use of computers can have an effect not only on children's physical health, but also shape children's social interactions and development. Children and especially young boys use computers mainly to play games (Subrahmanyam et al., 2000). Also, the dominant paradigm in computer interaction assumes one user on one computer at any time (Inkpen et al., 1999). However, several researches (Benford et al., 2000; Inkpen et al., 1995; Stewart et al., 1998; Stanton et al. 2001) show that children enjoy and benefit from collaboration with other children when using computers. As Inkpen et al. (1995) suggest, "children naturally gather in groups, especially to play games". Collaborative computer technologies are a thriving new field in children's computer use and focus on teaching new social skills to children.

QuiQui's Giant Bounce is intended for both home and classroom use in schools and day-care centers. The game can be played alone or with pairs as shown in Figs. 2 and 3.

Fig. 2. Oskari playing the second version of the flying game.

Fig. 3. Brothers playing together.

Our technology enables collaboration in the following ways: (1) children can physically play together, (2) children can help a friend by shouting or clapping hands when a friend is technically in charge of the game and (3) children can watch the gameplay, give instructions and give praise to others. Our collaborative gaming technology can be defined as a computer game that enables co-present users to collaborate on a shared computer with a single shared display and simultaneous sharing of input devices such as a web camera and a microphone. This is different from single display groupware (SDG) applications (Stewart et al., 1999) where each user has their own input device such as mouse or keyboard.

The drawback of most computer vision based collaborative experiments (Bobick et al., 1999; Stanton et al., 2001) is that they require specialized hardware, time-consuming installation or expensive appliances. We believe that using inexpensive web cameras, microphones and careful product design enable the use of perceptual interactions in every home. Crowley et al. (2000, p. 63) emphasize that the key to perceptual interaction is usability that determines the requirements for technological innovation. One of the main goals of our iterative usability research is to find out how well the implemented avatar responses meet the needs of children's physical and social development. On the basis of the test results and the children's opinions about the game, the technology and avatar responses will be adjusted to the children's way of playing in the specified story context.

2.4. Requirements for usability research methodology

We think that both child-centered and participatory usability evaluation and design methods are needed when designing physically interactive environments for children. In order to find a suitable design and usability evaluation method we analyzed the unique

characteristics of our software product to define requirements for the selection of the method. Based on our main focus on the study of the physical interaction itself and the social context where the computer game will be used, the usability evaluation method must enable us to:

- observe several children and provide them with an environment where they can feel comfortable even if they are video taped
- diminish problems caused by the child–adult relationship like authority and differences in language or knowledge
- provide a natural social context where children would play the game as if in a real world situation, and
- evaluate the learnability and physical appropriateness of the game tasks in the social context.

The analysis of the requirements implied that the usability observation and evaluation methods are more suitable approaches in our study instead of design-oriented methods like participatory design, low-tech prototyping (Druin et al., 1999) and cooperative inquiry (Druin, 1999). Even though several researchers (Alborzi et al., 2000; Montemayor et al., 2002) have successfully applied different design methods like low-tech and wizard-of-oz prototyping, and scenario walkthrougs needed for children and adults to co-design physical interactive environments we decided to focus on developing a usability evaluation method that addresses the special requirements set for the physically interactive computer game. However, we do believe that these participatory design methods are useful when designing the audio-visual story world for the game.

We felt that paper prototyping or any other low-tech prototyping methods are not suitable for our perceptual user interface design purposes, because the physical and vocal abilities needed to play the game are not something that can be designed or low-tech prototyped together with children. Although there has been some research on prototyping tools for perceptual user interfaces as of late (Sinha and Landay, 2001), we have not identified tools for prototyping such interfaces where an avatar mimics the user's movements. In our opinion, the best option for designing this kind of interaction is to implement a working prototype, test it in a real world like situation and reimplement it according to the test results.

Section 3 discusses the peer tutoring approach that was designed to make the usability evaluation of the physically and interactive computer game possible. The method provided us with information about how children play the game physically and vocally in an authentic social context with their age mates.

3. Peer tutoring approach

3.1. Usability and children

How should the concept of usability be defined when users are children? It seems that many researchers in the field of children's technologies agree that it is unclear whether the

standard definition of usability is suitable for evaluating children's technology. Hanna et al. (1999, p. 4) state that "the usability of a product is closely related to children's enjoyment of it." Druin et al. (1999, p. 67) suggest that if a tool is easy to learn and control, then children will quickly become immersed in the experience. This relates to Nielsen (1993, p. 27) statement that "learnability is in some sense the most fundamental usability attribute."

We think that one definition of the usability of a children's software application is that a child is able and willing to teach other children how to use it. This enables the expertise in using the software to propagate among children. In other words teachability is an important part of learnability. The more teachable and learnable the software is, the more potential users there will be. Using peer tutoring in usability evaluation allows us to evaluate teachability. In addition, it also provides us with crucial knowledge on how children use the system and how they communicate about the software: what kind of instructions and terms they use and where their language differs from that used by the designers of the software and other adults.

3.2. Definition of peer tutoring approach

A person entering the field of peer collaboration would be amazed by the number of theoretical and experimental frameworks in use. The extensive research of peer relations varies from a psychoanalytic approach to pedagogical practice to social psychology. Though many researchers in the field of HCI for children have acknowledged that collaborative technologies encourage peer learning and teaching (Cassell and Ryokai, 2001; Stewart et al., 1999, Stanton et al., 2001, Benford et al., 2000; Inkpen et al., 1999), peer tutoring has not been applied as such in a children's usability research before.

Multiple definitions for peer roles exist but they are not all consistent. Corsaro (1992) says peers are defined as a group of children who spend time together on a daily basis. In our research, peers are considered to be children of similar age and status, and belonging to a same classroom. Cross-age collaboration is not part of this research.

Peer tutoring is one type of peer collaboration. Damon and Phelps (1989a) define peer tutoring as an approach in which one child instructs another child in material in which the first child is an expert and the second is a novice. As Goodblad and Hirst (1989) put it, "peer tutoring is a system of instruction in which learners help each other and learn by teaching". The latter definition does not make any requirements of the expertise of the tutor but emphasises the fact that tutors normally benefit from teaching. In our study both the child who is the teacher (tutor) and the child to be taught (tutee) are not very far removed from competence due to the short learning period. The tutors in our study had a possibility to play the game two to three times according to their preference before they were asked to teach the other child.

The peer tutoring approach presented in peer collaboration literature cannot be applied as such in the context of usability testing. Thus, peer tutoring is based on several research fields (Fig. 4).

Usability testing can be a very unnatural experience for children if they are taken to try out an unknown computer product in an unfamiliar environment and tested by strangers while video cameras are recording. Even though this has proved to be a suitable solution for adults (Nielsen, 1993; Dumas and Redish, 1993), it is less fitting for usability testing

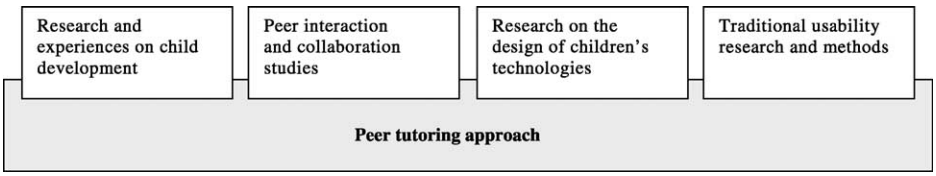

Fig. 4. Fields of research that have influenced the design of peer tutoring approach.

with children. Also, children have different skills from adults and those skills develop rapidly during childhood, which means that usability methods have to be flexible and modified to meet the children's needs. Thus, the research on child development must be taken into account when designing usability testing for children.

Hanna et al. (Hanna et al., 1997) have adapted traditional usability testing for children and give guidelines for the use of laboratory tests. Peer tutoring also has its foundation in more traditional usability practices like co-discovery learning (Dumas and Redish, 1993), thinking aloud (Nielsen, 1993; Boren and Ramey, 2000), co-participation (Wilson and Blostein, 1998) and paired usability testing (Wildman, 1995). The peer tutoring method differs from the co-discovery learning and co-participation methods in that it is not based on the idea that two participants work collaboratively on a given task but that the task of a tutor is teaching and the task of a tutee is to act out according to the tutor's instructions.

Traditional usability testing provides guidelines for organizing test sessions. However, even child-friendly usability testing in a laboratory (Hanna et al., 1999) does not solve problems like the lack of collaboration or natural social context, or problems due to child–adult relationships. We have applied the peer tutoring approach to usability testing with children because it is both collaborative and provides test results that are suitable for evaluating a physically and vocally interactive computer game for children. The benefits and challenges of using peer tutoring approach are real social context, equality in authority and knowledge, children taking an active role in usability testing, making communication easier and more lively, and getting authentic video material. These benefits and challenges are described in the next sections.

3.3. Real social context

The notion of the usability of interactive systems depends on the social context in which the interactive systems are used. An unfamiliar environment like a usability laboratory and observational technology (one-way mirrors, speakers and video cameras) accentuate children's awareness of being judged and observed despite reassurances that it is the software that is being tested. In the peer tutoring approach testing will be carried out in the natural environment of the children—homes, schools and day care centers with age mates they already know. Peers are usually so engaged with the software and with each other that they do not even notice the testing going on. Peer tutoring allows children to be more playful and exploratory and less goal-oriented when involved in a usability activity with a peer. Thus, peer collaboration provides us with an insight into the patterns of activity and the social gaming context.

If a collaborative product is tested, then it needs to be tested collaboratively. For example, as our intention is to distribute the game in Finnish day-care centers and elementary schools, we are interested in how children would use it in these environments and what kind of interaction takes place in a space where more than one child can use the game. We look forward to seeing whether collaboration between children emerges and how excited, motivated or bored they will be while playing the game, or observing someone else play it.

3.4. Equality in authority and knowledge

In a child–adult relationship, the differences in the levels of knowledge and authority affect the nature of the discourse between the child and the adult (Damon and Phelps, 1989a, p. 138). Thus, the challenge is to open up the communication and encourage children to be verbal and diminish the authority. For example, Druin (1999, p. 594) suggests that power structures can be changed by wearing informal clothing. Hanna et al. (1999, p. 11) recommend to 'establish a relationship with children when you first meet them by engaging in some small talk'. However, neither of these advice can always prevent power structures from emerging. Also shyness, the fear of giving wrong answers and children's need to please adults can affect the test results.

The best aspect of peer collaboration is that it provides tools for communication and equality. 'Children have certain advantages over adults in teaching peers. They may more easily understand the tutee's problems because they are cognitively closer' (Gaustad, 1993). As Damon and Phelps (1989a, p. 138) put it, "unlike adult–child instruction, in peer tutoring the expert party is not very far removed from the novice party in authority or knowledge; nor has the expert party any special claims to instructional competence. Such differences place tutee in a less passive role and make the tutee feel freer to express opinions, ask questions, and risk untested solutions." Rogoff (1990, p. 172) also implies that children are likely to treat the situation differently if they are in charge of it rather than being given a task by adults, which is the case in traditional usability testing.

One drawback of peer tutoring is that tutees, often labeled as less capable than tutors, tend to resist being tutored by their age mates (Gaustad, 1993). This can be prevented by not using tutors that are much more skillful than the tutees in playing the game. For example, all tutors in our formal tests were new to the game and only had a few minutes of play practice. Thus, the tutors were not vastly more experienced players than their tutees.

3.5. Children taking an active role in usability testing

The first experience children have with a new system is that of learning, which is not generally emphasized in HCI research. Peer tutoring provides us with information about the learnability of game tasks and what kind of instructions children use when teaching one another. In practice this means that one child teaches the other how to play the game and thus takes an active role as an instructor in a usability testing situation. Our study is based on a guided peer tutoring approach, meaning that researchers are still in the same space to help and guide the tutor–tutee interaction but only if necessary. Since children are

more active in the testing situation, researchers can stay more remote and adult intervention is diminished.

Damon and Phelps (1989a) also mention that children can manage their own interactions, invent their own problem-solving procedures, and discover their own solutions. Peer tutoring approaches that focus on peer collaboration to solve a problem are especially effective in fostering creativity, experimentation and problem-solving skills (Damon and Phelps, 1989b, p. 151 that are highly beneficial in testing a new kind of technology.

3.6. Making communication easier and more lively

The communication in traditional usability testing is likely to be asymmetrical between adults and young children. One reason why peer tutoring works is the fact that peers speak in a more similar manner than do adults and children. Also, children are more relaxed to communicate with adults when there is a peer in the test space. The interaction between tutor and tutee is more balanced and more lively, which is important in usability research that aims at studying the use of a product in a real social context. The tutor–tutee communication is highly valuable when analysing how well children have learned the required skills needed in the game, how they perceive the interaction and how much and what type of instructions are suitable for children.

Using peer collaboration is also useful because when children communicate with peers they pay careful attention not only to the partner's utterances, but to his or her intended meaning as well. Children are also sensitive to non-verbal cues offered by another child to show that they do not understand something (Goodblad and Hirst 1989). One of the problems in adult–child interaction is timing. Garvey (1986) suggests that adults tend to think that it is their responsibility to maintain a temporally coherent and constant pace in the conversation. Adults also tend to answer more quickly and appear to think that it is their duty to efficiently exchange turns in conversation. In talk engagements with a peer, the child must begin to assume a greater share of that responsibility.

Young children cannot be forced to 'think aloud'. The peer tutoring approach encourages children to verbalize their thoughts naturally and spontaneously. Even though Boren and Ramey (2000) strongly criticize the current practice of thinking aloud protocols in usability research, they do not consider child related issues.

Especially younger children need help in teaching the other child, which requires some adult intervention. Since our test situation is not completely adult-free, it is important to make sure that the researchers behave informally and make space for the children to interact with each other. The researcher–child communication requires some planning before the testing. In our study we decided to allow time for praise when children performed well and give instructions only when the children seemed to be frustrated or did not understand what to do. Researchers also used concepts that are more natural or imaginative to children like 'the eye of the computer' (web camera) or 'magic square' (the taped square on the floor).

3.7. Authentic video material

We believe that in evaluating usability of perceptual user interfaces, the recorded video material is crucial as both a memory aid and a tool for analyzing the movements children use to physically or vocally control the computer. For example, in evaluating QuiQui's Giant Bounce, we needed to observe several children to ensure that the selected movements be physically appropriate for as many children as possible. Since we needed to find out the ways in which children find it natural to use the game, videotaping is a key element in storing the physical action in test situations. Also, we have to make sure that all children in our target group are able to master the selected conceptual movements. The more natural the children's behavior in the video recording situation is, the more useful the video material is for design purposes. The peer tutoring approach provides a pleasant setting for children to act out physically and vocally in. Also, unlike Druin (1999, p. 594) we have not found the use of video cameras restricting or affecting the children's behavior.

4. Applying peer tutoring in practice

This study involved two usability tests using two versions of the flying game task. The contextual movements used in the first flying game prototype were flapping both hands up and down for flying upwards and flapping only one hand for banking to the side of the flapping hand. To get QuiQui to fall towards the ground, the child needs to stay still. In the second version of the game, to fly sideways, the child flaps both hands and leans his or her body to the desired direction.

In the first test, the first prototype of the flying game task was evaluated. Our main interest was to find out what the most natural way for children to control a flying avatar was. After redesigning the flying game task, the second test was held. In addition to these formal test sessions, several informal testing situations including 12 children were organized during the redesign of the flying game task. The usability tests took place at the Lycée franco-finlandais d'Helsinki in Helsinki, Finland. The school is bilingual—both French and Finnish are used in teaching. As the school consists of a preschool day care center, an elementary school and a high-school, we were able get children from the whole target group to participate in tests without having to change location. The test was held in a small but otherwise normal classroom.

4.1. Participants

The peer tutoring approach involves at least one tutor, a tutee, and an adult researcher called the *interactor* who is guiding the collaboration. The interactor does not operate the video camera or take notes, but teaches the tutors, guides the testing situation and asks questions during the peer collaboration efforts.

The participants included 28 children (12 in the first test and 16 in the second) between the ages of five to nine. Both boys and girls participated in the study, but due to the small number of children, gender-specific issues are not analyzed in this paper. The children were chosen by their teachers and none of the children who participated in the first session

took part in the second session. Also, children who had played the game before were not accepted to participate in the testing. The children's parents were asked for permission for their children's participation in the tests. Parents also filled in a questionnaire with their children to provide us with background information on the children. The background questionnaire revealed that 75% of the children had a computer at home and 46% had a video game console at home. Only 14% of the children did not have either a computer or a video game console at home. Most children were familiar with computers and had played computer games or used children's CD-ROMs. The children used computers in average 80 min per week according to their parents' estimates.

In the first test session, the children were divided into four groups of three children. In the second test, four groups of four children were used. The children in a group belonged to the same classroom, so they were familiar with working and communicating with each other. The test sessions lasted between 1 and 2 h, depending on how long the children wanted to explore the game freely. The children were not forced to play the game any more times than they really wanted to. The times for single flying games for all children were recorded to evaluate how much the change in the contextual movements affected the playing times.

4.2. Two-on-one and each one teach one approaches

The test setup consisted of a web camera, a laptop computer with a built-in microphone, and a video camera for recording the action for further analysis. The video recording was used to capture the child–child communication and the movements the children used to control the avatar. A 'magic square' was taped on the floor to help the children to stay in sight of the web camera and to allow them to focus more on the game instead of having to stare at the video capture window while playing. This prepositioning is done because the children should see their hands in a video capture window at all times during gameplay, whether they extended them fully to their side or whether they were on top of their head. The magic square also helped position the video camera as we could better estimate where the children would move in the space.

The testing was conducted with one group at a time. Two different peer tutoring approaches were used, depending on the number of tutors. In the first test session, *two-on-one tutoring* was used to prevent the child from being left alone with the researchers during the test. The second test was based on *one-on-one tutoring* where a child spends some time alone with the researchers but only after he or she has tried out the game first with his or her tutor. The latter approach is also based on the *each one teach one method*, where each tutee acts as a tutor for the next child, and thus transfers his or her experiences and knowledge about the game to the next child. Both the *two-on-one* and *each one teach one* approaches have a similar introduction phase. When the children enter the test room, the researchers introduce themselves and the interactor takes the lead and asks the children to approach the computer.

In *two-on-one tutoring*, two children from a group of three are paired to participate in the tutor training phase. The third child (tutee) stays in the classroom with the rest of the class and is asked to join in later during the test session. The test setup is introduced to the tutors by telling that 'the computer's eye (pointing simultaneously the web camera) sees

you and you should be standing inside the magic square so the computer is able to see you well'. The interactor then briefly explains the idea of the game and shows what the contextual movements are by flapping hands. After the introduction, the other child is guided to sit next to the interactor to follow the other tutor's playing. The game is started and the child plays the game while the interactor and the other child give further instructions and praise. After a round of playing the game it is the other child's turn. When the tutors have played the game a couple of times, the interactor then asks the tutors if they could teach their tutee how to play the game. The tutors ask the tutee to enter the test space and starts teaching the tutee how to play the game and shows the contextual movements to the tutee. The interactor starts the game and asks tutors to give further instructions while the tutee plays the game.

The second test was based on the *each one teach one* approach, where four groups of four children of the same age went through the each one teach one procedure. In the *each one teach one* approach, an interactor teaches the first child in a group how to play the game and allows the child to practice the game by playing a few times. The other children in the group wait in another room. The interactor then asks the child if he or she could be a tutor and teach the tutee how to play the game. The tutor then goes to fetch the tutee in the test space. After one round of gameplay, the tutor is asked to exit the test space and the tutee then has the opportunity to try to play the game alone. When a tutee has played the game two more times, the interactor asks if the tutee wants to be a tutor and so forth. Thus a group of four children consists of three tutors and four tutees.

4.3. Question asking protocol and guidance

The preschool aged children sometimes have problems in telling verbally what the other child should do or what happens in the game. If the younger children did not spontaneously teach the tutee, the interactor asked the tutors some extra questions like 'where should you be when playing the game'. Normally, this opened up the communication and the children started to explain enthusiastically "come here, come here! (went standing in the magic square), and do something (flapped their hands), and you have to hit the clouds!' Especially for 5–6 year olds this 'question asking from the tutors' proved to be a suitable option in helping the tutors when they got stuck in their teaching. We feel that with the peer tutoring method, it is important that the researchers not teach or instruct the tutees but ask questions from the tutors if the teaching situation requires adult intervention.

The *question asking protocol* is valuable as it combines both interviewing the tutor and provides help for the tutor in the teaching situation. We think that it is a good option in gathering the 'likes' and 'dislikes' from young children. The question-asking protocol is simple and effective. Two categories of questions were used:

- Questions that help tutors to teach a tutee. When the tutor is teaching a tutee, the interactor does not give any instructions to tutee but asks a game or activity related question from tutor when tutor seems to need help, or memory aid or the tutee is having serious trouble in playing the game.
- Comment related questions. When the tutor comments on an event in the game like 'you turned into pink!', the interactor asks a comment related question, for

example, 'what makes you turn pink?', and the tutor replies 'that pink jumping apple with big teeth'.

When answering a question, the tutor provides game related information based on the tutor's own experiences and observations using language similar to the tutee's. We also believe that question asking in peer tutoring provides more spontaneous and honest answers than formal interviewing because children do not necessary feel that they have to please adults. The tutors' answers give good insight into how the tutors have understood the interaction in the game. The problem with the question-asking protocol is that the tutors tend to direct the answers to the interactor rather than the tutee. We have noted, however, that a tutee listens to the tutor's answers carefully and uses the information in these answers to understand the game better. Additionally, tutees pass the knowledge obtained from the tutor's answers to their own teaching and answers.

In contrast to cooperative inquiry (Druin, 1999), guided peer tutoring does not restrict the children that are tutoring others from getting involved in the action and thus limit the communication between the children. However, we too think that adult intervention in the peer collaboration approach is problematic. It is difficult to be assertively passive while guiding the interaction. It is also hard not to get carried away when the children are very enthusiastic or to cheer them up when they are bored. The interactor in the test situation should not interfere with or lead the collaboration between the children any more than is actually required.

5. Results

5.1. Promoting natural communication

Peer tutoring promotes communication and natural interaction style in the test situation. It is related to asking the user to think aloud, a method commonly used in usability testing. However, constantly thinking aloud to the interactor can be difficult for children, especially if they are using an application with a heavy cognitive load such as a physically interactive game which requires constant monitoring of the game events. Boren and Ramey (2000) warn that any outside influence, any comment or prompt from the researcher turns subsequent verbalizations into those that require additional cognitive processing beyond that required for the task. We have also noted that when asking questions from a child playing the game, he or she either does not answer at all or stops playing for a while when answering.

The benefits of peer tutoring approach in communication are:

- Children need not be asked to think aloud—the teaching process provides communication that comes about naturally.
- The cognitive load is split—the tutees can concentrate on the task while tutors handle the communication by teaching the tutee or by answering the interactor's questions.
- Tutor helps researchers to see the situation through the eyes of the other child.

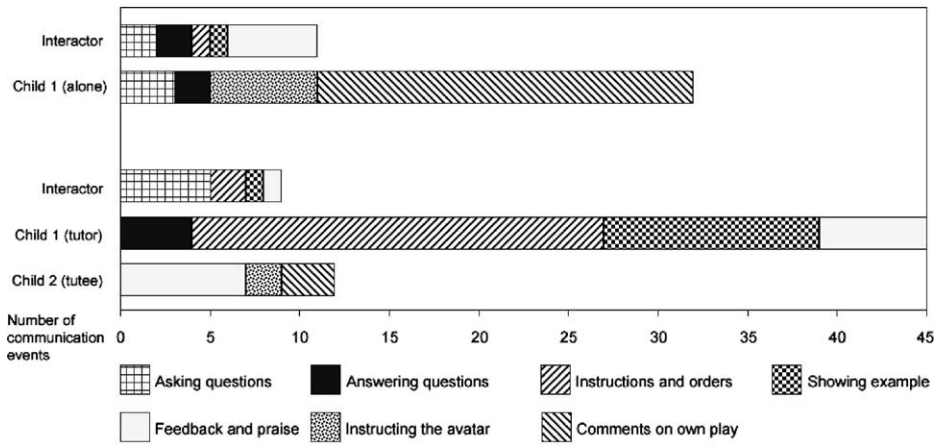

Fig. 5. Number of communication events during the two phases of each one teaches one session for 6 year old boys.

- Children are active and communicate lively. This result is based on counting the communication events, like phrases and single comments, during the test session and placing them in the following categories (Fig. 5): (1) asking questions, (2) answering questions, (3) giving instructions and orders, (4) showing an example, (5) giving feedback and praise, (6) instructing the avatar, and (7) commenting own play. In addition to these categories, the children also modeled the tutors' physical examples and the tutors usually emphasized their instructions by pointing at the objects on the screen they were talking about. Fig. 5 shows one example of how many times children communicated and what type of communication events took place.

  When analysing the communication in the *each one teach one* situations, the results show that:
- Tutors are very eager to give instructions and orders and show examples to their tutees.
- Tutees tend to either instruct the avatar or comment on their own play.
- In each case the change from the role of the tutee to the role of the tutor affected the communication. The tutees tend to be quiet during the teaching process and not ask a lot of questions from the tutor or the interactor. When the tutees became tutors they were more lively, talkative and made more eye contact with the people in the test space. The difference in communication was especially noticeable with quiet children.

The children used two kinds of instructions during the test. When tutors were watching their tutee play the game, instructions were directed to the tutee rather than the avatar. For example, one tutor said "Drop down, you will soon get the cloud (watching that the tutee does what is told). Good!" When children gave instructions during teaching, tutees followed the instructions carefully. The language used differs from that used by the interactor even though the children adopted similar concepts, like the 'magic square' used by the researchers while introducing the game. The children invented their own instructions when explaining the interaction to others like 'drop drop drop!!!' when the other child should not have moved for a while. Children also gave praise to each other and

Fig. 6. Rasmus showing Oskar how to fly.

instructed others in a friendly manner. When playing alone, most children commented on their own play and instructed the avatar spontaneously, which reveals what the child attempts to do and helps understand his or her mental model of the game. Even though this behaviour is useful in analysing the usability of the game, we do not believe that it is due to the peer tutoring approach but the children's natural behaviour when playing a computer game.

Fig. 7. Rasmus giving further instructions to Oskar.

Fig. 8. Claudia telling Jemina to step into the magic square.

5.2. Teaching patterns and modeling

According to our experience children only teach things that they like, understand and feel are important. Various teaching situations can be seen in Figs. 6–9. Our analysis of the children's teaching showed that a nearly similar pattern of teaching emerged in all groups. The pattern was (1) 'Go to the magic square' (children pointed the square taped on the floor), (2) 'You can see yourself here' (pointing the video capture window), (3) 'You have to flap your hands like this' (children showed an example by flapping their hands), (4)

Fig. 9. Claudia showing Jemina how to fly sideways.

‘There are clouds that you have to make rain. And when you have all these (pointing at rain drop icons) you have won’. It seems that children feel that a physical location in a room and the video capture window are very important in the game. Most children were very amazed about the fact that they could see themselves in the video capture window. They pointed to the window and looked at each other or the researchers and said ‘see!’.

Analysing teaching patterns is useful as it provides the researcher with insight into how children behave and categorize information, and more importantly, what they all find important with the product. The challenge for further developing the peer tutoring approach will be the research and study of learning theories and social interaction theories to better understand how teaching patterns emerge and how tutees actually learn from their peers. Also it is interesting to see how much the interactor’s teaching procedure affects the tutor–tutee interactions. One possible tool for analyzing the actual teaching–learning situation later on is Vygotsky’s theory about the zone of proximal development (Vygotsky, 1978) which he defined as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers”.

It was amazing how much observing and being shown an example affects the other child’s play. In *two-on-one tutoring*, the second tutor managed to control the character as if he or she had already played it before. This finding emphasizes how well children learn by observing others perform tasks. Modeling is an important part in child–child tutoring. All children showed examples and pointed at objects on the screen they were talking about. When children model, they offer each other a behaviour model for imitation. When showing how to control the avatar, children tend to first show only how to fly upwards, but when they watched their tutee actually play the game, they gave further instructions on how to control QuiQui and what the different objects in the game world meant. This finding can be used to design and incorporate a help system in the game based on video material demonstrating the actual use of the game rather than using auditive or static visual feedback. We believe that the peer tutoring approach is a good method for finding design solutions for the help system of the computer application that is being evaluated.

5.3. Capturing usability problems

The key to a successful usability evaluation of a perceptual user interface is authentic video material. Using the peer tutoring approach provided us with video material where children behaved naturally and did not seem to be distracted by the video camera at all. The communication between the tutor and the tutees, and the movements children used to control the game were analyzed using the video tapes in order to find usability flaws in the game.

The sideways flight control in the first prototype was problematic, and obviously frustrating for the children. The frustration manifested itself in the following ways: (1) non-task related activity and speech like wandering around the room and asking ‘where is the chess board?’, (2) calling the avatar names like ‘that QuiQui is a strange fellow, flying wherever’, (3) commenting and questioning the functionality directly like ‘I just cannot

make it fly in that direction', and (4) trying to control the avatar in ways not anticipated by the designers.

Several hours of video material were analyzed to find out what was wrong with the conceptual movements of the first version of the flying game task. We saw in the videotapes that the children leaned to the side they wanted the QuiQui character to fly to when they could not control the game in the way they expected. Due to this observation the contextual movements were changed so that to fly sideways, the child flaps both hands and leans his or her body to the desired direction. This change was thus based on how children actually tried to control the sideways flight when they did not manage to use only one hand for flying. When using an improved version of the flying game task, the children gave less frustrating comments and did not try to control the avatar in any other way than they were taught. To verify that the new version was more controllable, the playing times of the two tests were compared. The results show that the children spent on average 34% less time playing the new version, which is a significant improvement in the game controls. Also, some preschool aged children managed to spend less than one minute in collecting all eight clouds, a difficult task to accomplish even for adults.

The question asking and children's natural communication reveal the conceptual models children have of the game. Especially the problems in the user interface design came to our attention. We noted also that the visual feedback and the mapping between some of the game elements and the children's conceptual model of these elements requires further design. For example, not all children understood that the blue clouds were 'full of water' and the white clouds had already rained, even though a rain animation was played along with a rainy sound. Because the distance between the child and the computer screen is approximately 1.5 m, the visualization of the game world must be considered more carefully in further designs of the game.

5.4. Emphasis on test setup

If the method will be applied in the different usability evaluation context the following guidelines may help researchers to carry out the testing. The relationships of the participants are relevant for the peer tutoring approach to function. In the first test, one boy acted aggressively against his fellow male tutee, which lead to the interactor taking a more authoritative role and raising her voice to prevent the children from hurting each other. This kind of behaviour is in stark contrast to what the peer tutoring approach requires to be beneficial. Due to this incident, we propose that same sex participants should have a friend like relationship in their everyday life. Also, motivation is higher when teaching and playing with a friend. When the tutor and the tutee are not of the same gender, they tend to behave in a more respectful manner against one another.

Having several children in the test space worked out well since the children were relaxed and focused on observing, and participated in giving instructions. Even though the children were able to wait for their own turn quite well at the separate space, it is advisable to cut the waiting time down by having a properly designed schedule. Depending on the number of researchers participating the study the children could be in separate room waiting for their turn but this requires some adult taking care of them. Other good solution

is to let the other children to wait in their classroom and get the tutor to pick up the next child.

If children are located quite near the product to be evaluated we would recommend to use the each one teach one approach. We also recommend to locate the children so that the tutor cannot operate the computer and take over which is one of the challenges of using the peer tutoring method. In our case the tutor was sitting next to the interactor and gave instructions but did not tried to go on and play the came.

5.5. Challenges and drawbacks of the method

Even though the results and experiences of using peer tutoring are promising and it has proved to be effective in detecting usability flaws of the computer game, there are several drawbacks that we want to discuss to give a more balanced view on the method. As with many other usability evaluation methods the peer tutoring approach requires a lot of work in organizing the test sessions and analyzing the video material. The research team using the peer tutoring approach must be multidisciplinary since the approach requires expertise in the field of usability research and especially on working with children to be beneficial—and as we all know these people are not necessarily easy to find.

As mentioned earlier we think that a true participatory and child-centered development of children's physically interactive computer products would require both a suitable design and usability evaluation methods. The experiences got from the test sessions made us to think that the peer tutoring approach should be coupled with some participatory design method to become even more useful for the overall game design.

The analysis of the test sessions showed that all tutors taught their tutees correctly. This finding implies that the research challenge is not to find out what things children can teach to each other but what they leave unsaid and why. Therefore the method needs to be finetuned to get children to verbalize also their thoughts about design solutions that they do not fully understand or like. One possible approach might be to use a visual walktrough or to combine group play and interviewing to the test setup.

The usability evaluation of a physically interactive user interface is not an easy task since it requires both the study of the physical appropriateness and the analysis of the usability of the actual audio-visual user interface. We feel that the method was very useful in defining the physical interaction suitable for children in the defined target group but we hoped that it could have helped us understand children's mental models about the game world better than it did. This is why we aim to further study how the capturing of the usability problems of the actual game logic and the audio-visual material could be made easier. One of our future task is also to look deeper into the communication and learning patterns to gain more knowledge about the phenomena mentioned earlier.

6. Conclusions and future work

We believe that the peer tutoring approach is a valid and efficient method for evaluating the usability of a perceptual user interface with children, even though it requires time, careful planning of logistics and willingness to work with children. We also suggest that it

might be a good solution for testing more traditional software products. Additionally, we propose that the method could be beneficial in the children's computer education. The peer tutoring approach reveals both good and bad design solutions, as well as teaches the researchers how children could use the product in an actual social context. Peer tutoring is a good solution when:

- The software application to be tested is designed for social settings like schools, where it is beneficial that children can teach other children how to use the software
- A collaboration enabling computer product is being evaluated
- The software application is based on continuous interaction that has a heavy cognitive load and thus, the thinking aloud and interacting with the interactor is difficult for children
- Adult–child communication and power structures are sidetracking the test situation
- Children's natural behaviour and communication is crucial
- Usability problems need to be found out by analysing communication as well as observing behaviour
- Children's physical and spatial abilities need to be verbalized
- The design of a help system is based on children's views and experiences on the computer product
- A video recording is needed but it should not distract the users.

We found the *each one teach one* approach to be especially beneficial as it provided both paired and individual testing at the same time. The approach also works well when many teaching–learning events are needed. The benefit of the *two-on-one* approach is that children seemed to be a bit more relaxed and the role of the interactor was more in balance. Also, the *two-on-one* approach revealed how much children actually learn by observing other children perform a task.

Being inspired by the results and experiences obtained from the usability tests, we plan to continue with working and further experimenting with the peer tutoring approach. As the results of this study are still preliminary, more detailed guidelines for applying peer tutoring in practice have yet to be published. This research has raised a number of issues associated with age appropriateness, interactor's role, children's satisfaction in being tutors and tutees, as well as other peer tutoring approaches like peer collaboration and cross-age tutoring that we aim to study in the future.

Acknowledgements

First of all, we would like to thank all the children who participated in our usability tests, especially the students of Lycée franco-finlandais d'Helsinki. We are deeply grateful to our partners in The Young Finland Association, Faculty of Sport and Health Sciences at the University of Jyväskylä and 'Children are Telling Group' at the National Research and Development Centre for Welfare and Health (Stakes). We would like to acknowledge the contributions and advice given by many lecturers

and course instructors at the University of Tampere, Helsinki University of Technology, and the University of Art and Design Helsinki, UIAH. Currently, the work is supported by graduate schools UCIT and HeCSE. Financial support has also given by YLE (Finnish Broadcasting Company), the Alfred Kordelin Foundation, and the University of Arts and Design Helsinki UIAH. To all parties concerned, our deepest thanks.

References

Alborzi, H., Druin, A., Montemayor, J., Sherman, L., Taxen, G., Best, J., Hammer, J., Kruskal, A., Lal, A., Plaisant Schwenn, T., Sumida, L., Wagner, R., Hendler, J., 2000. Designing StoryRooms: interactive storytelling spaces for children. Proceedings of Designing Interactive Systems (DIS).

Benford, S., Bederson, B.B., Akesson, K., Bayon, V., Druin, A., Hansson, P., Hourcade, J.P., Ingram, R., Neale, H., O'Malley, C., Simsarian, K., Stanton, D., Sundblad, Y., Taxén, G., 2000. Designing storytelling technologies to encourage collaboration between young children. CHI 2000, ACM Conference on Human Factors in Computing Systems, CHI Letters 2 (1), 556–563.

Bobick, A., Intille, S., Davis, J., Baird, F., Pinhanez, C., Campbell, L., Ivanov, Y., Schutte, A., Wilson, A., 1999. The KidsRoom: a perceptually-based interactive and immersive story environment. Presence: Teleoperators and Virtual Environments 8 (4), 367–391.August.

Boren, M.T., Ramey, J., 2000. Thinking aloud: reconciling theory and practice. IEEE Transactions on Professional Communication 43 (3), 261–278.September.

Cassell, J., Ryokai, K., 2001. Making space for voice: technologies to support children's fantasy and storytelling. Personal Technologies 5 (3), 203–224.

Corsaro, W.A., 1992. Interpretive reproduction in children's peer cultures. Social Psychology Quarterly 55, 160–177.

Crowley, J.L., Coutaz, J., Bérard, F., 2000. Perceptual user interfaces: things that see. Communications of the ACM 43 (3), 54–64.March.

Damon, W., Phelps, E., 1989a. Critical distinctions among three approaches. In: Webb, N. M., (Ed.), Peer Interaction, Problem-solving, and Cognition: Multidisciplinary Perspectives, Pergamon Press, New York, pp. 9–19.

Damon, W., Phelps, E., 1989b. Strategic users of peer learning in children's education. In: Berndt, T., Ladd, G. (Eds.), Peer Relationships in Child Development, Wiley, New York, pp. 135–157.

D'Hooge, H., Goldsmith, M., 2001. Game design principles for the Intel® Play™ Me2Cam* virtual game system. Intel Technology Journal Q4, 1–9.

Druin, A., 1999. Cooperative Inquiry: Developing New Technologies for Children with Children, Proceedings of CHI'99, ACM Press, Pittsburgh, PA, pp. 592–599.

Druin, A., Bederson, B., Boltman, A., Miura, A., Knotts-Callahan, D., Blatt, M., 1999. Children as our technology design partners. In: Druin, A., (Ed.), The Design of Children's Technology, Morgan Kaufmann, San Francisco, pp. 52–72.

Dumas, J.S., Redish, J.C., 1993. A Practical Guide to Usability Testing, Ablex, Norwood, NJ.

Garvey, C., 1986. In: Mueller, E.C., Cooper, C.R. (Eds.), Process and Outcome in Peer Relationships, Academic Press, Orlando, pp. 329–345.

Gaustad, J., 1993. Peer and cross-age tutoring. ERIC Digest, 79.

Goodblad, S., Hirst, B., 1989. Peer Tutoring, a Guide to Learning and Teaching, Kogan Page, London.

Hanna, L., Risden, K., Alexander, K., 1997. Guidelines for usability testing with children. Interactions 4 (5), 9–14.

Hanna, L., Risden, K., Czerwinski, M., Alexander, K., 1999. The role of usability research in designing children's computer products. In: Druin, A., (Ed.), The Design of Children's Technology, Morgan Kaufmann, San Francisco, pp. 4–26.

Inkpen, K., Gribble, S., Booth, K.S., Klawe, M., 1995. Give and take: children collaborating on one computer. Proceedings of CHI'95: Human Factors in Computing Systems, ACM Press, 258–259.

Inkpen, K., Ho-Ching, W., Kuederle, O., Scott, S., Shoemaker, G., 1999. This is fun! We're all best friends and we're all playing: supporting children's synchronous collaboration. Proceedings of Computer Supported Collaborative Learning (CSCL) '99, Stanford, CA December, 252–259.

Montemayor, J., Druin, A., Farber, A., Sims, S., Churaman, W., D'Amour, A., 2002. Physical programming: designing tools for children to create physical interactive environments. Proceedings of CHI2002 Conference, Minneapolis, MN, 299–306.

Nielsen, J., 1993. Usability Engineering, Academic Press, Boston.

Pinhanez, C.S., Wilson, A.D., Davis, J.W., Bobick, A.F., Intille, S., Blumberg, B., Johnson, M.P., 2000. Physically interactive story environments. IBM Systems Journal 39 (3–4), 438–455.

Rogoff, B., 1990. Apprenticeship in Thinking: Cognitive Development in Social Context, Oxford University Press, Oxford.

Sengupta, K., Wong, H., Kumar, P., 2000. Computer vision games using a cheap ($<$ 100$) webcam. Proceedings of Sixth International Conference on Control, Automation, Robotics and Vision (ICARCV'2000), Singapore December.

Sinha, A.K., Landay, J.A., 2001. Visually prototyping perceptual user interfaces through multimodal storyboarding. Proceedings of Workshop on Perceptual User Interfaces, Orlando, Florida November.

Stanton, D., Bayon, V., Neale, H., Ghali, A., Benford, S., Cobb, S., Ingram, R., O'Malley, C., Wilson, J., Pridmore, T., 2001. Classroom collaboration in the design of tangible interfaces for storytelling. Proceedings of Human Factors in Computing Systems (CHI2001), ACM Press 2001, 482–489.

Stewart, J., Raybourn, E.M., Bederson, B., Druin, A., 1998. When two hands are better than one: enhancing collaboration using single display groupware. Proceeding of CHI98, 287–288.

Stewart, J., Bederson, B.B., Druin, A., 1999. Single display groupware: a model for co-present collaboration. Proceedings of CHI99, 286–293.

Subrahmanyam, K., Kraut, R.E., Greenfield, P.M., Gross, E.F., 2000. The impact of home computer use on children activities and development. Children and Computer Technology 10 (2) Fall/winter.

Topping, K., 1988. The Peer Tutoring Handbook: Promoting Co-operative Learning, Billing and Sons Limited, Worcester, Great Britain.

Vygotsky, L.S., 1978. Mind in Society, Harvard University Press, Cambridge, MA.

Wildman, D., 1995. Getting the most from paired-user testing. Interactions July, 21–27.

Wilson, C., Blostein, J., 1998. Pros and cons of co-participation in usability studies. Usability Interface 4 (4).