Head Up Games : on the design, creation and evaluation of interactive outdoor games for children

DOI:
10.6100/IR754836

Document status and date:
Published: 01/01/2013

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Head Up Games

On the design, creation and evaluation of interactive outdoor games for children

IRIS SOUTE

Head Up Games

On the design, creation and evaluation
of interactive outdoor games for children

Head Up Games - On the design, creation and evaluation of interactive outdoor games for children, by Iris Soute

This research was supported by the European Community under the Information Society Technologies (IST) programme of the 6th Framework Programme (FP6-2004-IST-4) Project PASION (Psychologically Augmented Social Interaction Over Networks).

A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 978-90-386-3393-0

Cover design and layout: Saskia Renders
Photography: Iris Soute
Printed by: Gildeprint Drukkerijen

Head Up Games

On the design, creation and evaluation
of interactive outdoor games for children

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van
de rector magnificus, prof.dr.ir. C.J. van Duijn, voor
een commissie aangewezen door het College van
Promoties in het openbaar te verdedigen op
dinsdag 18 juni 2013 om 16.00 uur

door

Iris Anna Catharina Soute

geboren te Stein

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. P. Markopoulos

en

prof.dr.ir. J.B.O.S. Martens

Table of Contents

List of Head Up Games

Chapter 1

Introduction

Spring 2005, students are starting their design case projects as part of the USI program. I am one of them and have joined a group that has not yet formed a definitive goal for the project. We have chosen our general direction though: we are going to design "something for children".

This thesis is inspired by the design that ensued from that design case: an outdoor, interactive game for children.

First, let's have a closer look at the situation in 2005. Many researchers are addressing children as a user group within the context of computer use because there are many concerns. There are concerns about obesity as a result of a sedentary lifestyle (Ebbeling et al., 2002). Furthermore, concerns are raised about the solitary nature of computer games having a negative impact on the social development of children (Subrahmanyam et al., 2001). Finally, there are concerns about the influence of violence in games on children's behavior (Anderson and Bushman, 2001).

Not only concerns are raised. Opportunities are identified as well: new technology can provide opportunities in facilitating play between normal developing children and disabled children (Brederode et al., 2005); it can provide novel learning experiences (Sluis et al., 2004) and it can motivate to be more active (Mueller et al., 2003).

A common theme in the research mentioned above is that it includes games and play. In the former examples in a rather negative light, as it is argued that computer use has a negative impact on children's development and lifestyle. In the latter examples, we see cases of the use of novel technologies to wrap an end goal (learning, facilitating interaction between children) in a game. Evidently, children are linked with games and, on a more abstract level, fun; the common idea being that fun is a catalyst for engaging and motivating children.

The situation described above was the situation in which we started defining our design case. The brief of our project was to design a game that supported social interaction between children. In contrast to earlier designs that pursued a deeper goal by wrapping it into a fun game, we decided to design for fun as our main goal, fun being the main motivator for children to engage in the game. We argued that, though learning games of course have their merit, creating a game that provides an opportunity for children to play together, offers many benefits by itself: in child development literature it is argued that play is of great importance for the development of children (Scarlett et al., 2004).

This resulted in the design of *Camelot* (Verhaegh et al., 2006). Reflecting on *Camelot* it became clear that it was the first example of a new, unexplored, genre of games. *Camelot* showed the possibilities of merging elements of traditional play with digital play, resulting in a fun game for children.

THEORIES OF PLAY AND DEFINING GAMES

Before elaborating on the research challenges that we identified after *Camelot*, in this section we will first review existing theories of play and elaborate on the definitions of play and games. Note that this is not an exhaustive review, only the main theories will be discussed that hold relevance to this thesis.

First we should note that "play" is a rather ambiguous word. Consider the following two statements: "Let's go outside and play tag!" and "Shall we play house? You are dad and I am mum". Though in both statements the same expression is used, completely different activities are referred to. Play, and the equally ambiguous term: games, have been studied in many different fields (Sutton-Smith, 1997), e.g., by historians, psychologists, sociologists, anthropologists, mathematicians etc. As we are interested in children's play and games, we will limit ourselves here to theories of play that relate to children. Often theories of play are linked with child-development theories. It is generally thought that play holds a significant role in a child's development (Yawkey and Pellegrini, 1984), though theorists are not always in agreement as to in what manner.

The earliest theories of play, the classic theories, mostly relate to biological/physical aspects of play. For example, the surplus energy theory of play assumes that children play to release excess energy (Spencer, 1873). In contrast, the relaxation theory of play argues that play is a means for children to restore their energy levels (Patrick, 1916).

In the twentieth century new theories were formed, based on insights gained in developmental psychology. Probably the most widely known theory is that of Piaget (1962), who

Age	*Stage*
Age 0 - 2	Practice play - is characterized by the repetitive practice of simple motor skills, e.g., clapping of the hands.
Age 2 - 6	Make-believe and symbolic play - children imitate actions of others, and engage in fantasy play. They will use objects to symbolize something different than the intended function (e.g., pretend that a broom is a horse)
Age 6 - 12	Games with rules - children's reasoning becomes more logical and play becomes more structured. Games played with peers are prefered, such as tag.

Table 1.1 Piaget's stages of play, with approximate ages

argued that children's cognitive development progresses through several distinct stages. Piaget thought that play held a significant role in this process and linked the development of play to the child's cognitive development (see Table 1.1 for the three stages of play). Furthermore, Piaget argued that play's development reflected a child's developing capacity to think in an abstract (or symbolic) manner.

Similar to Piaget, Vygotsky (1976) believed that children actively learned through play. However, whereas Piaget believed that a child mostly independently progressed in his cognitive development, Vygotsky stressed the importance of social interaction to mediate cognitive development. Furthermore, he theorized that learning occurred in the zone of proximal development: the distance between a child's ability to perform a task under guidance of, for example, a parent and a child's ability to accomplish the task independently. Besides mediation of a parent, Vygotsky argued that play, too, could create such a zone, and could thus foster development. Finally, Vygotsky assumed that children acquired social rules by engaging in make-believe play.

Though many researchers today build on these theories, there has also been critique; for example, Piaget's theory has been said to leave untouched the fact that play occurs in many settings and contexts, and that these settings and contexts affect the way children play and develop (Berk, 2002). Contemporary theories that do address this are referred to as cultural-ecological theories; an example of research in this area is Pellegrini's work on children's playground behavior during school recess (Pellegrini, 1995). He argued that behavior during recess has impact on classroom behavior and on the social and cognitive development of children. Furthermore, Rogoff (2003) draws on the theories of Piaget and Vygotsky, but emphasizes the influence of culture on child development and also on play. For example, she argues that individualistic cultures are more likely to feature competition in play, in contrast to cultures that value interdependency.

A recent survey of research in educational and behavioral domains on children's play shows that 'play' is still a prominent topic for research (Cheng and Johnson, 2010); most studies surveyed originated in the educational field and research focuses on the implications of findings in practice, not on developing new theories on play. The majority of researchers operationalized play according to their own particular research purposes. Most research seems to be directed at the role of play in a child's cognitive development, by using play (or games) as a scaffold for learning activities in school. For example, children's development theories are applied in interface design for educational games, for an overview see Revelle (2013). Typically these efforts build on Piaget's and Vygotsky's theories.

The impact of physical activity play on child development receives far less attention; though evidence is still scarce, Pellegrini and Smith (1998) hypothesize that physical play too has an important role in a child's social and cognitive development. Also, recently people have started worrying that the decline in outdoor play is affecting children's development (Gray, 2011; Louv, 2008; Rivkin, 2012).

We highlighted relevant theories of play. For a more elaborate overview, the reader is referred to for example Mellou (1994) or Scarlett et al. (2004).

Though we have reviewed now theories of play, we have not yet established what a "game" exactly constitutes, or have provided a definition of what "play" is. First, play and games are inevitably intertwined and much ambiguity exists. Actually, Sutton-Smith (1997) addresses this in his book "the Ambiguity of Play" stating that "We all play occasionally, and we all know what playing feels like. But when it comes to making theoretical statements about what play is, we fall into silliness. There is little agreement among us, and much ambiguity" (Sutton-Smith, 1997). In an attempt to bring coherence to this ambiguous field, instead of presenting clearly outlined definitions, he presents the seven rhetorics of play, building on existing theories and placing these in a broader context. The rhetoric that is relevant to children's games is the rhetoric of play as progress, i.e. the notion that children adapt and develop through play (Sutton-Smith, 1997, Chapters 2, 3).

Indeed, since play and games are so diverse, it is probably impossible to find a definition that covers all play and games. Salen and Zimmerman (2003) have reviewed several definitions of play to come to an understanding of what a game actually constitutes. Definitions of play-scholars of different fields have been examined, see Salen and Zimmerman (2003, p. 79), such as of Huizinga, Caillois and Costikyan, and common elements have been identified. From this Salen and Zimmerman form the following definition of a game: " a game is a system in which players engage in an artificial conflict, defined by rules, that result in a quantifiable outcome."

In contrast, instead of trying to derive a definition of play by listing its distinct game elements, Scarlett et al. (2004) and also Sutton-Smith (1997) define play as what children themselves see as play: having fun, being outdoors, being with friends, choosing freely, playing games, not working, pretending. In this thesis we have adopted a similar manner of looking at play and games; our notion of Head Up Games is inspired by how children naturally play outdoor games.

It is not the intention of this thesis to add new insights to the existing theories of play, or form a new definition of games. The type of play we are addressing is well described in existing theories of play and as such Head Up Games may arguably contribute to children's development - though it is not our goal to find evidence for this. Instead we, without confirming or disconfirming any of the existing theories, are approaching play and games from a design perspective: the notion of Head Up Games was formed as a counter movement to the, at the start of the work of this thesis, available pervasive games for children. These games predominantly rely on screen-based interaction, which, as we argue, interferes with play patterns found in traditional games. Arguably, such interaction could diminish some of play's positive effects on children's development.

RESEARCH CHALLENGES

Camelot was positively received in the research community, and we started reflecting on the reasons of its success. *Camelot* appeared to be different than anything available in the market or anything developed in the research field: the game design focused on fun as it primary purpose, the technology did not require focusing on a screen and the resulting game appealed to the children, prompting them to be physically active and socially interactive. Inspired by this we formed our notion of Head Up Games. Head Up Games are:

outdoor, co-located, multiplayer pervasive games that encourage social interaction, physical activity and support adaptable rules, creating a fun experience

All of the work in this thesis is pivoted on that notion. In researching Head Up Games we took a multidisciplinary approach, addressing the work from three perspectives: design, creation and evaluation. In the context of our aim for Head Up Games each perspective poses specific challenges. Note that we make a distinction between *design* and *creation*. Often these terms are used intermixed; in this context we define *design* as the process of designing the games, and *creation* as the actually building (i.e. hardware and software development) of the games.

Throughout this thesis we often refer to "children", which can be rather broadly interpreted from babies to teenagers. We focus on children from 7 to 10 years old. The main reason for doing this is that the games we aim to design can be described as "games with rules". From child development literature (e.g., Piaget (1962), Scarlett et al. (2004)), we know that children start playing games with rules from the age of 7. Below that age children's play mostly constitutes of pretend play or imaginary play; they are not yet engaging in games with rules, making this age group unsuitable for playing Head Up Games.

DESIGN

Designing outdoor games for children using ubiquitous technology poses some specific challenges to the design process. Adopting a User Centered Design (UCD) approach, we look for ways to involve children in the design process. In the field of Interaction Design and Children (IDC) this is a topic often researched, and many methods are available. Our challenge lies in finding the right method in the context of game design.

Game Design literature (e.g., Salen and Zimmerman, 2003) advises a similar process as UCD, i.e. an iterative design process. It is recommended to start testing early, and test with low-fi prototypes. However, in the context of Head Up Games this proved to be a challenge; many interaction styles used in Head Up Games cannot be approximated using low-fi prototypes. The pace of a Head Up Game is often high and the player experience is unquestionably altered when interaction styles are implemented as mock-ups in a Wizard of Oz setup.

Main Challenges - Design

Our challenges in designing Head Up Games concern a) how to meaningfully involve children in a game design process and b) determining the most appropriate way of deploying prototypes in such a process.

CREATION

Throughout the work of this thesis multiple Head Up Games have been created. However, each time the hardware and software for the games were developed from the ground up. Though common functionalities could be identified, often the technology of previous games could not be reused for new games. Building prototypes for Head Up Games brings about issues that are common when creating embedded hardware and software, e.g., difficulties integrating the hardware of several technologies. Furthermore, typical for Head Up Games is that the footprint of eventual devices should be kept small and robust as the devices are intended for outdoor use. Finally, we note that the actual implementation in hardware and software is often pushed to the end of a design process. Consequently, technological issues and the impact on the overall game experience are identified very late. Therefore, we argue that a prototyping tool is necessary to be able to deploy interactive technology in an earlier stage of the design process.

The target user-group of such a prototyping tool are game and/or interaction designers, people who are typically not trained as embedded hardware and software engineers. For these designers to adopt such a tool, we venture into the field of End-User Programming. More importantly, we are interested in how the use of a prototyping tool would influence the design process of game designers.

Main Challenges - Creation

The main challenges identified in this perspective are a) defining a set of interaction styles appropriate for Head Up Games, and implementing these in a prototyping platform; and b) the uptake of the platform by game and/or interaction designers and how their design process is influenced by this.

EVALUATION

We want Head Up Games to resemble traditional outdoor games, like tag and hide-and-seek. As such, the play behaviors we aim at are physical activity and social interaction, generating a fun activity. In the field of Human-Computer Interaction (HCI) observations methods are commonly used to quantify behaviors. However, these methods mostly assume that only one subject is involved, and more importantly, that that subject remains relatively stationary throughout the observation. In contrast, the games we intend to design involve multiple children at the same time, and children to run around all the time.

Furthermore, the HCI community is more and more focusing on the user experience (UX) of designs. The typical approach is to identify individual constructs that contribute to the overall experience. These constructs, e.g., immersion, enjoyment, flow can be quantified using questionnaires or other types of surveys, see for example Poels et al. (2007), though it is argued that using post-play methods, the experience cannot be properly measured (Stenros et al., 2012). If children are involved, the challenge becomes greater: the level of literacy of children can interfere with their ability to fill in a questionnaire, and furthermore, children tend to be very positive about the product under evaluation anyhow (Markopoulos et al., 2008). An example

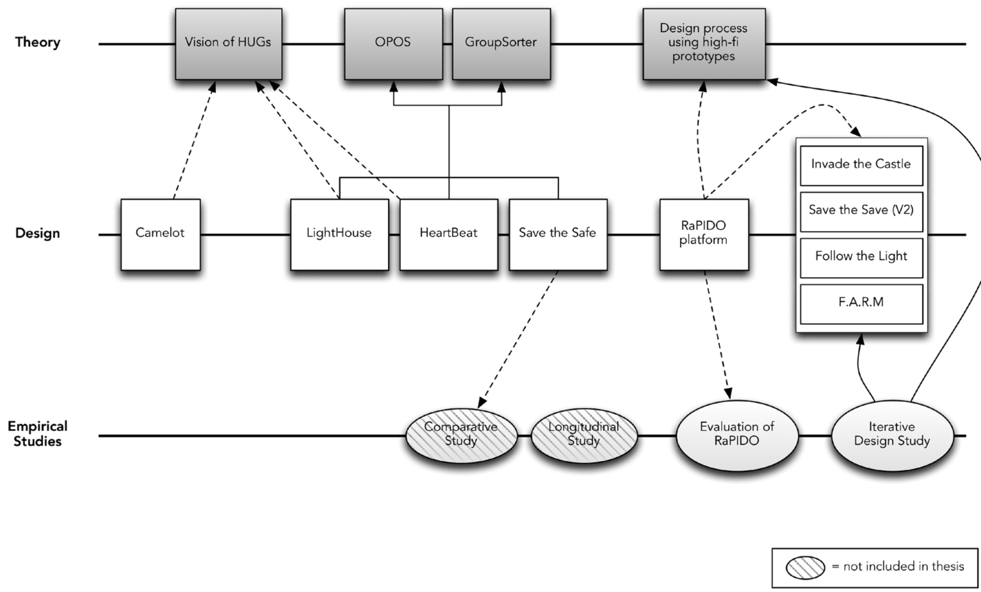


Figure 1.1 Overview of the work, showing relationships (modeled after Mackay and Fayard (1997))

of a toolkit that takes these challenges into account is the Fun Toolkit (Read, 2008).

Concluding, there are methods available to observe and survey children, however we need to establish whether or not they are suitable in the context of outdoor gaming.

Main Challenges - Evaluation

Our challenges with regard to evaluating Head Up Games is to find appropriate methods for measuring play behaviors and fun, while doing this in a way that is appropriate for many children at once in an outdoor context.

METHODS

Each perspective described above requires its own dedicated methods for approaching the challenges. The overall approach for this thesis is a Research through Design (Zimmerman et al., 2007) approach: knowledge about designing Head Up Games is built up by actually designing and creating them. However, in this approach we combined theory development, design and empirical studies. More specifically, the following methods and studies were used (note that not all studies are reported in this thesis), see Figure 1.1. Note that each design included an evaluation with children too, though these are not separately depicted in the figure:

- A vision document was put forward that captured the key aspects of Head Up Games, contrasting them to existing related work (Chapter 2).
- Several games were designed, implemented and evaluated in iterative design processes. All games were evaluated with children in context, i.e. in the locations were children normally play games. The reflective accounts of these design processes led to lessons learned about the design process of Head Up Games (Chapter 3)
- A case study approach was followed for researching the impact of using technology rich prototypes from an early stage in the design process (Chapter 4)
- An experimental comparison study for comparing the effect of a virtual and a physical game element on the game play (Soute et al., 2009).
- A longitudinal study for researching the adaptability of game rules (Soute and Sturm, 2011).

OVERVIEW OF THE THESIS

The challenges identified in the previous section have been addressed in this thesis. First, details of our vision of Head Up Games can be found in Chapter 2. Next, in Chapter 3 we present lessons learned about the design process of Head Up Games, and we present a study in which we iteratively designed four Head Up Games, involving children in the process. We also reflect on the use of high fidelity prototypes, i.e. functioning, interactive prototypes instead of mock-ups, from an early stage in the design process. During the design process, a rapid prototyping tool was used; its conception and creation are reported in Chapter 4. Moreover, also reported in Chapter 4, is a study that followed the design process of two designers while creating interactive games for children. In the study we focused on how the use of the rapid prototyping tool influenced their design process. In Chapter 5 we reflect on two methods for measuring the play experience in Head Up Games.

Though we report the work from the three different perspectives of design, creation and evaluation, in actuality the work has been carried out in a different order. Figure 1.1 presents the order of this work and how individual steps informed each other.

The games designed served as examples in more than one of our perspectives. To avoid repetition, we introduce each game at the point were it first appears in the thesis. Subsequently, if it appears again in the text, the game description is not repeated. Instead, the reader is referred back to the initial description, by way of a page reference. A list of the games presented in this thesis can be found on page 7.

REFERENCES

Anderson, C. A., and Bushman, B. J. (2001). Effects of Violent Video Games on Aggressive Behavior, Aggressive Cognition, Aggressive Affect, Physiological Arousal, and Prosocial Behavior: A Meta-Analytic Review of the Scientific Literature. Psychological Science, 12(5), 353–359.

Berk, L. E. (2002). Child Development (6th ed.). Allyn & Bacon.

Brederode, B., Markopoulos, P., Gielen, M., Vermeeren, A., and Ridder, H. de. (2005). *pOwerball*: the design of a novel mixed-reality game for children with mixed abilities. In Proc. of the 2005 Conf. on Interaction Design and Children (pp. 32–39). ACM, Boulder, Colorado.

Cheng, M.-F., & Johnson, J. E. (2010). Research on Children's Play: Analysis of Developmental and Early Education Journals from 2005 to 2007. Early Childhood Education Journal, 37(4), 249–259. doi:10.1007/s10643-009-0347-7

Gray, P. (2011). The decline of play and the rise of psychopathology in children and adolescents. American Journal of Play, 3(4), 443–463.

Ebbeling, C. B., Pawlak, D. B., and Ludwig, D. S. (2002). Childhood obesity: public-health crisis, common sense cure. The Lancet, 360(9331), 473–482.

Louv, R. (2008). Last child in the woods: Saving our children from nature-deficit disorder. Algonquin Books.

Mackay, W. E., and Fayard, A.-L. (1997). HCI, natural science and design: a framework for triangulation across disciplines. In Proc. of the 2nd Conf. on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (pp. 223–234).

Markopoulos, P., Read, J. C., Macfarlane, S., and Höysniemi, J. (2008). Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers. Morgan Kaufmann.

Mellou, E. (1994). Play Theories: A contemporary review. Early Child Development and Care, 102(1), 91–100. doi:10.1080/0300443941020107

Mueller, F., Agamanolis, S., and Picard, R. (2003). Exertion interfaces: sports over a distance for social bonding and fun. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 561–568). ACM, Ft. Lauderdale, Florida, USA.

Patrick, G. T. W. (1916). The psychology of relaxation. Houghton Mifflin.

Pellegrini, A. D. (1995). School recess and playground behavior: Educational and developmental roles. State Univ of New York Pr.

Pellegrini, A. D., & Smith, P. K. (1998). Physical Activity Play: The Nature and Function of a Neglected Aspect of Play. Child Development, 69(3), 577–598. doi:10.1111/j.1467-8624.1998.tb06226.x

Piaget, J. (1962). Play, dreams and imitation (Vol. 24). New York: Norton. Retrieved from http://www.psych.utoronto.ca/users/peterson/Psy2302011/03Piaget.pdf

Poels, K., De Kort, Y., and Ijsselsteijn, W. (2007). "It is always a lot of fun!": exploring dimensions of digital game experience using focus group methodology. In Proc. of the 2007 Conf. on Future Play (pp. 83–89). ACM, New York, NY, USA.

Read, J. C. (2008). Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. Cogn. Technol. Work, 10(2), 119–128.

Revelle, G. (2013). Applying Developmental Theory and Research to the Creation of Educational Games. New Directions for Child and Adolescent Development, 2013(139), 31–40. doi:10.1002/cad.20029

Rivkin, M. S. (2012). Children's outdoor play, an endangered activity. In D. P. Fromberg & D. Bergen (Eds.), Play from Birth to Twelve: Contexts, Perspectives, and Meanings (2nd ed.). Routledge.

Rogoff, B. (2003). The cultural nature of human development. Oxford [UK]; New York: Oxford University Press.

Salen, K., and Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals (illustrated edition.). The MIT Press.

Scarlett, W. G., Naudeau, S. C., Salonius-Pasternak, D., and Ponte, I. C. (2004). Children's Play (1st ed.). Sage Publications, Inc.

Sluis, R. J. W., Weevers, I., Van Schijndel, C., Kolos-Mazuryk, L., Fitrianie, S., and Martens, J. (2004). Read-It: five-to-seven-year-old children learn to read in a tabletop environment (pp. 73–80).

Soute, I., Kaptein, M., and Markopoulos, P. (2009). Evaluating outdoor play for children: virtual vs. tangible game objects in pervasive games. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 250–253). ACM, Como, Italy.

Soute, I., and Sturm, J. (2011). Towards a Setup for Longitudinal Evaluation of Player Experience with Children. Presented at the 2nd Int. Workshop on Evaluating Player Experience in Games (EPEX11), Bordeaux, France.

Spencer, H. (1873). The principles of psychology (Vol. 2). New York: D. Appleton.

Stenros, J., Waern, A., and Montola, M. (2012). Studying the Elusive Experience in Pervasive Games. Simulation & Gaming, 43(3), 339–355.

Subrahmanyam, K., Greenfield, P., Kraut, R., and Gross, E. (2001). The impact of computer use on children's and adolescents' development. Journal of Applied Developmental Psychology, 22(1), 7–30.

Sutton-Smith, B. (1997). The Ambiguity of Play. Harvard University Press.

Verhaegh, J., Soute, I., Kessels, A., and Markopoulos, P. (2006). On the design of *Camelot*, an outdoor game for children. In Proc. of the 2006 Conf. on Interaction Design and Children (pp. 9–16). ACM, Tampere, Finland.

Vygotsky, L. (1976). Play and its role in the mental development of the child. In J. S. Bruner, A. Jolly, & K. Sylva (Eds.), Play: Its Role in Development and Evolution. New York: Basic Books.

Yawkey, T. D., & Pellegrini, A. D. (1984). Child's play: Developmental and applied. Lawrence Erlbaum.

Chapter 2

Introducing Head Up Games Combining the best of both worlds by merging traditional and digital play

This chapter is based on:
Soute, I., Markopoulos, P., and Magielse, R. Head Up Games: combining the best of both worlds by merging traditional and digital play. Personal and Ubiquitous Computing 14, 5 (2009), 435–444

ABSTRACT

Current pervasive games are mostly location-aware applications, played on handheld computing devices. Considering pervasive games for children, we argue that the interaction paradigm that existing games support limits essential aspects of outdoor play like spontaneous social interaction, physical movement, and rich face-to-face communication. We present a new genre of games called "Head Up Games" to underline that they liberate players from facing down to attend to screen-based interactions. This chapter discusses characteristics of Head Up Games and relates them to existing genres of pervasive games. We present three examples of Head Up Games and reflect on the design and evaluation of these games.

INTRODUCTION

The availability of context-sensing and augmented-reality technologies has led to the emergence of pervasive games: "[a] genre in which traditional, real-world games are augmented with computing functionality, or, depending on the perspective, purely virtual computer entertainment is brought back to the real world" (Magerkurth et al., 2005). As a concept 'pervasive gaming' has been linked to numerous research fields (Nieuwdorp, 2007), such as ambient intelligence, augmented/mixed reality, mobile computing, location-aware computing, virtual reality and smart toys. The research reported below concerns specifically outdoor pervasive games for children.

In an outdoor setting, 'pervasive' is typically assumed to equal 'location-based' and numerous such games are reported in literature. Well-known examples are *Can You See Me Now* (S. Benford et al., 2003; Steve Benford et al., 2006) and *Catchbob!* (Nova et al., 2006). *Can You See Me Now* was developed partly as an artwork and partly as a research vehicle to test location-based technologies outside the safe environment of a laboratory; it is a catch game involving online players and players on the street who traverse an actual city chasing the (virtually present) online players. *Catchbob!* was designed to explore location-awareness in the context of mobile collaboration. In groups of three players have to find a virtual object by enclosing that object in a triangle formed by the participants positions in real space.

The games mentioned and many others of this type are played with a GPS and/or Wi-Fi enabled device, and use a small display (PDA or mobile phone) to show location and other game-related information. By targeting adult players in outdoor environments this genre of games has brought about a novel, non conventional, pastime and experience. For children however, outdoor play is a natural and traditional occupation and one that is essential for them: it has often been argued that play provides ample learning opportunities and is beneficial for a child's development (Scarlett et al., 2004; Vygotsky, 1976).

Current research literature is sparse when it comes to outdoor pervasive games for children. Well-known examples are *Savannah* (Steve Benford et al., 2005), and *Ambient Wood* (Rogers et al., 2004). In *Ambient Wood*, children take a field trip in a wood that is augmented with mobile and fixed devices providing contextually-relevant information. Each pair of children carries a PDA and a probe. In *Savannah*, children are equipped with PDAs with Wi-Fi and GPS. A virtual savannah is overlaid on a school field, and children have to cooperate as "lions" to hunt the savannah for virtual prey.

Both applications are successful demonstrations of a playful approach to learning. Unlike the focus of the research presented here, these applications involve a structured and externally motivated educational activity rather than a game that is designed to be played more than once and purely for fun.

Looking at pervasive games we note the reliance either on location-bound infrastructure, e.g., placing the devices in the forest for *Ambient Wood*, or accessing virtual content indexed by physical location, e.g., using GPS for location awareness. However, GPS errors can adversely influence the game play (Steve Benford et al., 2005), and using location awareness does not automatically lead to better game performance (Nova et al., 2006). Without ruling out location awareness, it appears that it would be fruitful to explore alternative mechanisms to support interaction between players and game content.

The brief overview above shows that the topic of pervasive games for children is still underexplored. Addressing this gap we introduce a new genre of pervasive games targeting children. These games are especially suited for late childhood (6-11) when children are able to follow game rules. The focus on outdoor play also brings about the need for interaction mechanisms more consistent with outdoor children's play. To emphasize the difference between pervasive games running on handheld devices which force players to "look down", the envisioned games are called "Head Up Games" (Soute and Markopoulos, 2007).

Head Up Games are a subset of pervasive games: they are played by co-located players in the physical world taking advantage of pervasive technology. In the following section we unpack the concept, compare it to related developments in the broader research field, and discuss experiences gained from the design, development, and evaluation of several Head Up Games for children. We conclude with reflections on this vision and directions for future research.

HEAD UP GAMES

In the "old" days, before the introduction of technology,

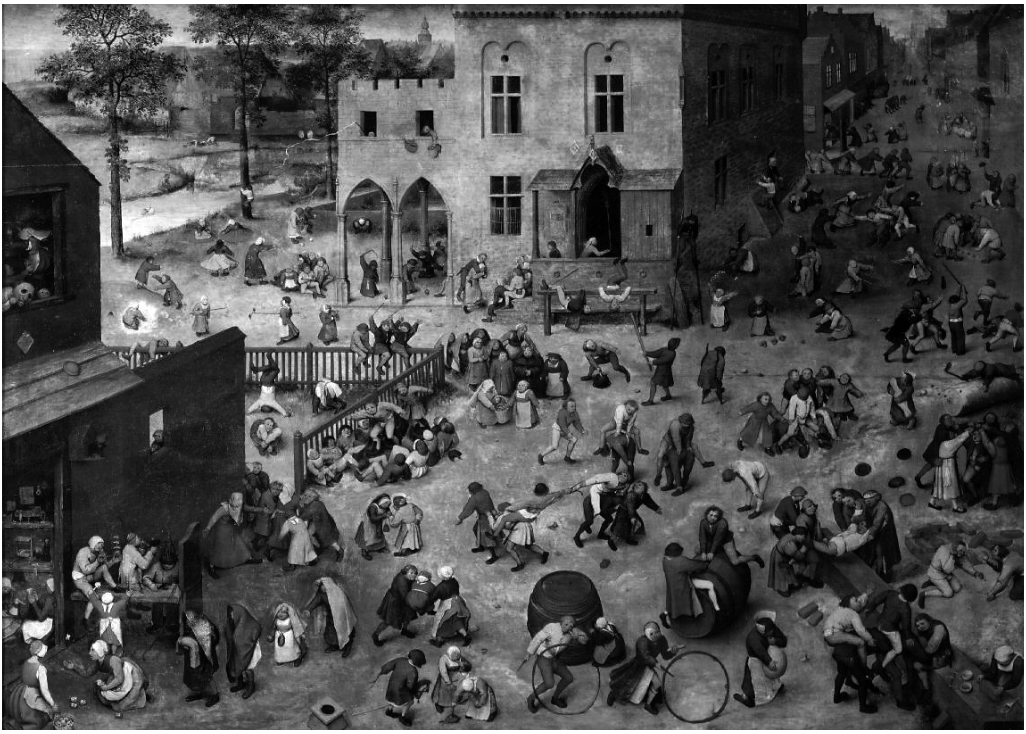


Figure 2.1 Children's Games by P. Brueghel, 16th century (Bruegel, 1560)

entertainment for children was fairly straightforward. A child played with toys, or with other children. A good example of the types of games played is nicely depicted in a painting of Brueghel (see Figure 2.1). Many of the games illustrated are still popular today, like tag, hide-and-seek, and ball games.

These games have a few characteristics in common: they require physical activity, and are played with few basic materials that children can easily take along (like a ball, a loop, or a skipping cord). As these games are played with multiple players in the same physical space, these games are rich in social interaction. Finally, the rules are often few and simple, appropriated and adapted by players. For the remainder of this thesis we will refer to this type of games as traditional (outdoor) games.

If Brueghel were to paint his picture again nowadays, it would probably show an abandoned playground. Compared to previous generations children spend less time playing out of doors, at least in industrialized societies (Clements, 2004; Gray, 2011). Children are spending an increasing amount of time indoors for less active and solitary pastimes, e.g., watching television, surfing the internet or playing games on consoles (Curtis, 2009; Turkle, 2011).

Child development theories acknowledge that play has an important role in a child's development (Scarlett et al., 2004; Vygotsky, 1976). For example, make-believe play helps form abstract thinking, and playing with other children leads to development of social skills. So, spending more of their time engaged in solitary activities can deprive children of developmentally valuable experiences. Some theorists even go as far as claiming that declining "backyard" play typical for previous generations, children are in danger of not developing morally (Scarlett et al., 2004) or link it to increased anxiety and depression and a reduced sense of personal control (Gray, 2011). Although it can be reasoned that internet provides a medium for forming and maintaining social relationships, concerns have been voiced about the depth and quality of these relationships and how they compare to offline social relationships (Cummings et al., 2002; Hellenga, 2002).

We argue that it is possible to combine the appeal of indoor digital games with the benefits of traditional outdoor play, thereby creating rich gaming experiences for children. Paraphrasing Marzano's maxim in "La Casa Prossima Futura" (Marzano, 2008) the vision of Head Up Games can be put forward as "the games of tomorrow will look more like the games of yesterday than the games of today". Our aim is to create:

outdoor, co-located, multiplayer pervasive games that encourage social interaction, stimulate physical activity and support adaptable rules, creating a fun experience.

Social Interaction. A game can provide the venue for a range of social interactions to unfold, e.g., competition, cooperation, negotiation, etc. While in many multi-player computer games social interaction is possible, it is typically limited to online chat or other forms of mediated communication. A position maintained by media theorists is the superiority of face-to-face communication over mediated interactions in its ability to convey non-verbal cues and, more fundamentally, the subtleties of verbal and non-verbal communication (Short et al., 1976; Turkle, 2011). Head Up Games, like traditional games, should take the advantage of the co-presence of players in the playing field, and should encourage a broader range of social interactions.

Physical activity. As an input to the design of *Camelot* (Verhaegh et al., 2006), observations and interviews regarding what makes outdoor play fun for children showed that children favor games that are rich in physical activity. Certainly, many traditional outdoor games feature physical activity as a component of the game.

Flexible and Adaptable Rules. Children engage in games with rules especially during late childhood (ages 6-11) (Manning, 2012). Traditional games often have flexible rules, e.g., a game of chase can be played two against two, but is equally playable in two teams of seven, or seven against one, which makes them playable in different contexts. Winning and losing rules do not have to be set in stone, but are negotiated: "once you are caught you are out", or, "once caught, you can be set free by a team mate". Specifically around the age of 10-11, children enjoy adapting game rules making rules increasingly complex and precise through a process of negotiation. Addressing the preferences and abilities of different children, the available resources, adhering to game rules, ensuring fairness in the play are important elements of their socialization process (Scarlett et al., 2004, p. 79).

Fun. Many pervasive games for children pursue an educational goal, e.g., Rogers et al. (2004), often resulting in a game with a narrative designed to transfer knowledge about a certain topic. Although these games can be a fun experience too, typically they can only be played once, since then the knowledge has been assimilated. We aim at creating games that are primarily aimed at having a fun experience, that children ideally will keep returning to, in a similar way that children keep playing games like tag and hide-and-seek.

With the above in mind, consequences with respect to technological choices for Head Up Games can be outlined. First, in contrast to other pervasive games Head Up Games should make minimal use of screen-based technology, since this can compete with the ability to interact directly with other players. A typical image for these games is of players attending to their devices 'head down' and attempting to observe and act within a virtual world through a handheld device. While interesting and engaging experiences are thus achieved, this form of interaction diminishes opportunities for direct face to face social interactions between players. Another drawback of a screen-based interaction with the game world is that it does not go together well with running around.

We place less emphasis on creating a virtual game world and meshing it with the physical world, focusing mostly on supporting the game play. Vygotsky (1976) has argued how make-believe play has a crucial developmental function to free thought from perception. The reliance of current game technology on explicit audio and visual effects competes with children's use of their own imagination. Instead of focusing on visual aspects, we aim to employ different interaction styles in the game and research alternative possibilities such as the use of sensors to sense for example motion, contact, or proximity.

Furthermore, whereas location-bound infrastructure or bulky technology restricts portability, traditional games can be played in any space that is appropriately large and safe;

all that is needed are the players, knowledge of the rules, and perhaps some simple game objects. An ideal to aim for is to assume as little as possible regarding the pervasive technology infrastructure required (e.g., do not assume GPS coverage or Wi-Fi connectivity).

HEAD UP GAMES AND RELATED CONCEPTS

Some of the characteristics of Head Up Games are also present in related genres of interactive applications and games.

Exertion interfaces (Mueller et al., 2007) are applications that motivate people to do sports activities where remote players compete in physically exhausting games. Activity monitoring and health coaching applications aim to motivate people to increase physical activity. The mobile phone application *Houston* (Consolvo et al., 2006) creates awareness of the user's own activity level and lets users share their activity information with friends. Sensing general activity level and using this within a game has been demonstrated in a series of games described as *Neat-O-Games* (Fujiki et al., 2007) and has even been applied in a commercial context for the Nintendo DS platform.

Whereas exertion interfaces have focused on adult players, research in Intelligent Playgrounds examines the use of intelligent objects and environments, to create a stimulating playground for children. Lund et al. (2005) present building blocks ('tangible tiles') that are reconfigurable into different games that detect and respond to the children's movements. Seitinger et al. (2006) describe how an interactive pathway impacts children's play patterns in outdoor playgrounds. They observed that many different play patterns emerged from a simple design that allowed for open-ended play. The *Interactive Slide* (Soler-Adillon and Parés, 2009) is a compelling application that is a mix between exertion interfaces and intelligent playgrounds aiming for co-located, collaborative and physically active play.

Bekker et al. (2008) present the concept of intelligent products for open-ended play. In open-ended play no fixed game structure is offered; instead the intelligent toys provide a setting where children are free to create their own games. Given their interest to support play and physical activity, the games considered by Bekker et al. bear many similarities with Head Up Games. However, a Head Up Game is not open-ended: open-ended play does not include game rules. From a player perspective the distinction is best described by Salen and Zimmerman when discussing the "Magic Circle" of play (Salen and Zimmerman, 2003, p.95) "…with a toy it may be difficult to say exactly when the play begins and ends. But with a game, the activity is richly formalized. The game has a beginning , a middle, and a quantifiable outcome at the end… Either the children are playing Tic-Tac-Toe or not." Head Up Games promote flexibility of game rules and players developing game rules of their own, assuming that at least a subset of the rules is implemented in technology.

Furthermore, we mention related work in the area of Social Games, i.e. games that have been specifically designed for enhancing social interaction. The game *pOwerball* (Brederode et al., 2005) is a mixed-reality tabletop game that resembles a pinball game. Two to four players have to collaborate to free imprisoned creatures. It was designed to encourage social interaction between children with and without a physical learning ability. Cheok et al. (2006) present *Capture the Flag*, a mixed-reality social game on smart phones. The game is played in teams of real-world and virtual-world players who communicate through text-messaging. In contrast to this example, Head Up Games are played in a co-located setting, which makes communication by way of game devices superfluous.

Outside the context of gaming, researchers in the field of multimodal and mobile interaction actively investigate head-up, or 'eyes-free', interaction for users on the move. These efforts focus on supporting users to interact with their PDA's or smartphones without the need to focus on the visual displays of these devices. Currently, the interface designs of many mobile devices predominantly rely on the visual modality. However, users on the move needing to interact with their devices cannot simply turn their visual attention to their devices, so other means of interacting with mobile devices while walking or driving are researched.

For example, Brewster et al. (2003) explore gesture-based interaction and audio feedback to enable users to use head movements for selecting audio items and they conclude that sound and gesture can significantly improve the usability.

Furthermore, Pielot et al. (2011) devised a system for pedestrian navigation using tactile feedback and compared that a regular (visual) system. The results suggest that using tactile feedback only will reduce the user's distraction, while combining both modalities (tactile and visual) will lead to a more efficient navigation performance. Finally, Robinson et al. (2009) propose *SweepShake*, a tool for discovery and information filtering of geo-tagged digital content in the space around the user. By making a sweeping gesture with a mobile device, the device will produce tactile feedback once digital content is found at the user's current location. In a study comparing *SweepShake* with a similar system using a visual interface, Robinson et al. conclude that users were able to find and select information with the tactile feedback only, though it took them more time compared to using the visual interface.

These three examples of 'eyes-free' show a focus on the replacement of visual GUI tasks with multimodal interaction techniques. Evaluations typically focus on comparing the accuracy and effectiveness of the novel interactions with the screen-based interactions. In contrast, our research in Head Up Games does not share this focus, as there are no visual GUI tasks to replace to begin with. Therefore, the evaluation methods used are of little value in Head Up Games evaluations. However, the context of use and styles of interaction in eyes-free interaction and Head Up Games are quite similar and insights in interaction styles in eyes-free interaction can inform the design of Head Up Games.

One can question whether a neologism is needed to describe the type of games we are designing. In the review of related genres of games we have shown how the defining characteristics of Head Up Games render them distinct from other genres, so at least the term has clear semantics. The distinct meaning though is not the primary reason for introducing the term; rather our aim is to take a clear position on the priorities designers and researchers should be setting when designing outdoor games for children.

EXAMPLES OF HEAD UP GAMES

This section discusses *Camelot*, *Save the Safe*, and *HeartBeat*, and reflects shortly on our experiences designing and evaluating these games. We will also reflect on how well we were able to adhere to our vision of Head Up Games.

CAMELOT

Camelot (Verhaegh et al., 2006), see textbox on next page, was created *before* we put forward our notion of Head Up Games. However, many design ideas for Head Up Games originated from *Camelot*, and as *Camelot* itself contains key characteristics of Head Up Games, we view *Camelot* as a Head Up Game.

Camelot involves several types of social interaction: competition, cooperation, discussion of tactics and team play. Evaluations revealed that fun was derived from the social interaction, the competition between the teams but also the suspense added by the unpredictable appearance of a ghost. This led us to consider the introduction of virtual elements in Head Up Games and to examine their role more systematically in follow-up Head Up Game designs. Furthermore, the nature of play supported by *Camelot* was well balanced. Intense physical activity alternated with periods of rest in such a way that *Camelot* did not resemble a dedicated sports application; something children appreciated during the evaluation sessions.

With *Camelot* we demonstrated the potential of supporting pervasive games using other technologies than screen technologies and how fairly 'simple' technologies can provide novel and enjoyable play experiences.

Figure 2.3 Left: Player acquiring resources. Right: building the castle at the castle construction site

CAMELOT

by Janneke Verhaegh, Iris Soute, Angelique Kessels

The aim of Camelot (Verhaegh et al., 2006) is to be the first team to complete their castle. To achieve this, each team is assigned a castle construction site (see Figure 2.3), where the castle must be built from several parts (castle towers, a drawbridge, a wall and a moat). To acquire these parts, players must gather different types of resources (water, brick or wood) in certain zones in the playground, by physically going to that zone, and staying in the zone for at least 10 seconds. The game is subdivided in four phases. In each phase a different part of the castle needs to be built, and each part requires a different combination of resources to be collected. Players can store a limited amount of resources at the castle construction site, for use in a later phase, but players are also allowed to trade resources.

To gather the resources players use collectors (see Figure 2.4). Each team has one collector per resource type. To get a resource, for example wood, a player takes the wood-collector and heads towards the resource zone (see again Figure 2.4). By placing the collector on top of the zone, acquiring of the resource starts. Acquisition takes approximately 10 seconds, then the player can retrieve the collector and bring it back to the castle site. Alternatively, the player can decide to gather a second piece of wood; in total 2 pieces of a resources can be 'loaded' on a collector. On the top of the collector are two leds, indicating how many resources are loaded on the collector.

At the castle site, a researcher notes what resources are brought in. If enough resources for a part of the castle have been gathered, players are allowed to build up the corresponding part of the castle.

The collectors and zones are implemented using PIC microcontrollers supporting the game logic, connected to infrared technology for communication between devices and LED lights to provide feedback to players. Each device functions as a stand-alone unit; hence there is no need for centralized computer control. The collectors weigh very little so children can easily carry them while running around. The zones are small and portable, so the game can be played anywhere.

Figure 2.4. Top: Resource zone. Bottom: collector

SAVE THE SAFE

Save the Safe (Soute and Markopoulos, 2009), see textbox on next page, was developed as a follow-up to the game *Stop the Bomb* (Hendrix et al., 2008), since the latter game had not been fully implemented.

In a final summative evaluation we compared two versions of the game. In one version the key was represented *virtually*: a player's belt vibrated if in possession of the key. In the other version the key was represented *physically*, by a ball.

Save the Safe is a high-paced game verging on a sports game, albeit one supported by pervasive technology; this was especially true when the ball was used to represent the key. Surveying participants' experiences showed how both versions of the games were more appealing to physically inclined players and, by the same token, less suitable for players who enjoy applying tactics in a game. Concerning social interaction: since the game was high paced, there was less time to discuss team tactics in depth. However, we did observe many yells and encouragements for other teammates.

With respect to using a virtual or a physical game object, it turns out that both approaches have their merit. However, we found it difficult to compare the two versions of the game as the resulting game dynamics are completely altered by the key being virtual or physical. Especially the visibility of the object seems to influence the game play: a physical game object is easily visible by all players, and it is straightforward for players how to act to transfer the key from one player to another. In contrast, the virtual game object, because of its invisibility, leads to higher suspense in the game, as players are mostly uncertain as to who has the key, and so a guessing game ensues of deciding who to pursue (or not) in an attempt to get the key.

HEARTBEAT

HeartBeat (Magielse and Markopoulos, 2009) was designed to explore the possibility of using biofeedback in Head Up Games.

Biofeedback is incorporated in the game using players' heart rate as input. Each player wears a heart rate sensor and during the game a player's heart rate is monitored. If the heart rate exceeds a preset value, the player's device

SAVE THE SAVE

Adaptation of Stop the Bomb by Koen Hendrix, Guo Yang, Dirk van de Mortel and Tim Tijs

Save the Safe revolves around obtaining (or guarding) a key to break into (or protect) a safe. At the start of the game the players are divided into a team of cops and a team of robbers; the cops win when they successfully guard the key from the robbers for the duration of the game. The robbers win the game when they steal the key from the cops and unlock the safe.

Each player wears a belt (see Figure 2.5). Mounted on the belts are a few LEDs and embedded in the belts are a vibration motor and a communication unit. At startup the units automatically form an ad hoc network. The communication unit is not meant for communication between players, but is used for determining distances between players, by measuring the signal strength of nearby belts.

At the start of the game the team roles are randomly assigned to the players. The belt shows the color of each player’s team, and the virtual key is initially given to one of the cops. Besides serving as an indication of the player’s team, the belts also indicate by vibration possession of the virtual key. If a player possesses the key, his belt starts vibrating while other players receive no such feedback. If another player approaches the key-bearer sufficiently close, the key is automatically transferred to this other player’s belt.

Figure 2.5 Above: the belt. Left: Players checking each others's team colors

starts broadcasting the heart rate to nearby devices of the opposing team, letting them know that opponents are near. So, besides running away from attackers, players can avoid being found either by hiding in bushes to stay out of sight physically or striving to keep down their heart rate to remain ‘hidden’ virtually. The tension between these tactics adds interest and suspense to the game play.

HEARTBEAT

by Remco Magielse

HeartBeat can be described as an adaptation of Capture the Flag that incorporates biofeedback in the game. Players are randomly divided into a defending team and an attacking team. One of the defenders is assigned a virtual treasure and, for the defending team to win, needs to remain untagged by an attacker for the duration of the game. The attacking team wins when they tag the defender with the virtual treasure and thereby capture the treasure. So, during the game the attackers need to hunt down defenders and tag them. Once tagged, a defender must join the attacking team. Defenders can protect the player with the treasure; when a defender has teamed up with the player with the treasure, both are protected against a single attacker.

Figure 2.6 HeartBeat devices showing team colors.

Each player carries a small game device (see Figure 2.6) and a heart rate monitor. If a player's heartbeat rises above 100 beats per minute, a signal is transmitted to nearby opponent players, and the opponent players' device starts to beep, indicating that an opponent is near. After tagging, a change of team is effected by docking the attacker device to that of the defender. Using the same interaction technique, the virtual treasure can be passed on between defenders.

The devices contained a PIC processor and an XBee module for communication between devices. Furthermore an interface was built to connect to an off-the-shelf heart rate monitor.

From the evaluation of *HeartBeat* we learned that compared to *Save the Safe, HeartBeat* is a more mixed-paced game, combining both running and hiding in a large area. Children related well to the idea of gaming outdoors in a technologically enhanced way. Heart-rate sensing was very novel for the children and seemed to have a positive influence on the game, although the evaluation results were not unambiguous.

REFLECTIONS

In this section we reflect on several aspects of the design, implementation and evaluation of the previous games.

First, we reflect upon the goals for Head Up Games and in how far we have been able to incorporate them in our designs. Second, we focus on lessons learned during the design process of Head Up Games, and how technology is interweaved in that process. Finally, we discuss issues that arose in the user evaluations.

CHARACTERISTICS OF HEAD UP GAMES

First an important word of caution is worthwhile to note: it is deceptively easy to design a Head Up Game that children appear to have fun with. This is especially the case when we evaluate the game experience in a single evaluation session. This is true for children's games in general; the evaluation of a game is typically (or it should be) set up as a pleasant activity in which children play a central role (Markopoulos et al., 2008). Achieving continued success and motivating children to return to a game is much harder. Children return to traditional games when they have a need for the physical activity and the social interaction that these games support. On the other hand, reasons to return to computer games include: to enjoy the evolving challenges and competition that these games provide, the variation on the audiovisual elements or the fantasy elements supported by a game narrative (Olson, 2010). In the reported examples of Head Up Games, the latter aspects that are traditionally found in gaming are still not sufficiently developed. To ensure long-term success, more effort needs to be spent on for example developing reward structures, leveling, and generally motivating children to return to the game.

The idea of encouraging social interaction has been a long-standing aim for research in games for children, prior to the emergence of Head Up Games (see survey of related work in earlier sections). Evaluations of these games have found, without exceptions, that they are successful in encouraging social interactions. However, the goal of social gaming has remained very loosely defined. It is necessary for researchers working in this field to become more specific about the forms of social interactions that games should encourage and perhaps about the developmental benefits that such games are expected to offer. Such specificity could help move this research field further from developing well motivated games to delivering concrete benefits to players. Hendrix et al. (2009) present an example of this more focused approach to supporting the development of social skills in children.

Encouraging physical activity was quite an easy task from a design perspective. Players could be motivated to be active in several ways, either directly setting them the challenge to escape or catch an opponent or, more subtly, by creating a tension between the need to run and the need to develop tactics that would allow a player to remain invisible (e.g., in *HeartBeat*). Whereas the effect during a game may be directly quantified by an observation scheme like OPOS (Bakker et al., 2008), see also Chapter 5 of this thesis, this is only part of the picture. Achieving the highest degree of physical activity would equate a game to sports, it could be played until the children are exhausted and then abandoned. Concrete benefits regarding a less sedentary lifestyle for children can only be evaluated in the long term. Games like the ones described above can have a positive effect if children choose to play these rather than be inactive in the playground or even playing these games as opposed to other handheld computer games.

In setting out to design and develop Head Up Games an important tenet was to avoid screen-based interaction. This choice has paid off, as we have shown that it is possible to design engaging game experiences without screen technology, but retrospectively this requirement seems somewhat extreme. Screens allow rich feedback and efficient information exchange so they could have a useful role in a Head Up Games as long as they do not command the player's sustained and concentrated attention. Rather than avoiding screens at all costs, the development of Head Up Games should proceed to explore novel interaction styles fitting outdoor play. The interactions shown in these games go beyond those traditionally used in augmented reality and tangible user interfaces like 'pick and drop', or 'selection' of virtual entities. Physical objects are used as containers of resources that can be passed from player to player by contact (*Camelot*), resources and game objects are separated into those that are movable by players and those that are bound to one location (*Camelot*), players can interact with each other by proximity (*HeartBeat*, *Save the Safe*), and may own properties unknown to themselves but made observable through the devices to other players (*Save the Safe*, *HeartBeat*). The list continues, but it is clear that developing novel and new interaction styles can spawn novel and exciting game experiences.

Camelot | p. 34

HeartBeat | p. 37
Save the Safe | p. 36

So far, the ambition to have flexible and adaptable rules has received little attention in our designs. As can be expected of research prototypes only part of the game rules were implemented in the technology. We did notice though, that in all the evaluations discussed and in the various phases of the design process children or even adult test participants would tend to modify those game rules that the technology did not enforce. This could be for a variety of reasons, for example to make the game more exciting, or to balance unequal abilities of players. There is nothing unusual about this behavior neither is it surprising that players tend to look for ways to cheat the technology once they know how it works (Consalvo, 2007; Salen and Zimmerman, 2003). Technology can provide an interesting contribution to this process of adaptation and appropriation, when changes in the implemented game rules can be effected by users and eventually maintained for future occasions or shared between players. The research challenge that emerges in the development of Head Up Games is the consideration of games as content created, shared and modified by communities of players.

GAME DESIGN AND TECHNOLOGY

As the emphasis is on outdoor play, an obvious way to design Head Up Games is by augmenting traditional outdoor games with technology. This augmentation can support different aims: to enforce rules and facilitate logistics or to provide new and engaging sensorial experiences through technology. The downside of this approach is that the added value of the technology is then rather limited: provided the games are already played by children, technology can be just 'sugaring' added on top of the traditional game but not an essential aspect of game play.

Another approach, which we followed with the design of *Camelot* and to an extent with the design of *Stop the Bomb* and *Save the Safe* is to start with a lo-fi prototype, play-test it fast and gradually introduce technology. This approach is consistent with traditional prototyping approaches for interaction design and with traditional game design methods (Salen and Zimmerman, 2003).

We discovered in practice that this approach can lead to problems. While the focus on fast iteration and on play-testing were found indeed beneficial to the designed games, we also experienced that first making a game that is truly playable in a low-tech version and only then introducing technology can bring about the problems regarding the augmentation of traditional games: the added value of the technology is marginal and design opportunities can be overlooked. In the design of *Heartbeat* we sought to introduce technology early on, also to explore the relevant design space. This design approach can lead to better integration and use of technology, making game designs that are not playable without the introduction of technology. Time-outs, biofeedback, random allocation of teams, proximity sensing, actions that are concealed from other players, balancing difficulty with randomness are some of the possibilities technology offers that could be poorly utilized if an already playable game is augmented.

Furthermore, technology offers the opportunity to add virtual game objects, characters or properties to the game that are not visible to all players but felt or heard by only a few. From the game evaluations we conclude that these elements can add to the suspense of the game.

We underline that it is not necessary (or even feasible) to capture all rules in technology. Many rules are automatically enforced by players themselves. For example, we considered to detect with technology who was tagged (initially, there was a rule in *Camelot* that required tagging, but was later dropped from the final version); however, we soon reconsidered this notion. Firstly, because the required technology would be complicated to implement, and secondly it turned out that for players themselves it is perfectly clear when one is tagged or not.

Designing game rules for outdoor play where players are co-located differs from designing computer games. The aforementioned tendency by players to make their own rules, extending and modifying existing ones, is a relevant design aspect. During play-testing the designer can deliberately provide only a small and loose set of rules. In trying to understand this rule-set players are bound to adapt them, fill-in gaps, resolve ambiguities and remove game-stoppers. Subsequently, the resulting set of new rules can be used in the next iteration of the development.

CONCLUSION

This chapter has summarized the development of a novel genre of outdoor pervasive games for children that are collectively described as Head Up Games. Head Up Games have some defining characteristics that set them apart from other pervasive games: they encourage physical activity while steering clear from being a sports game, they support rich social interaction face to face by co-located players while avoiding to transpose them into a virtual world or to focus their attention on game content and away from the other players, and they rely on simple adaptable rules that can be modified by players.

Three instructive examples of Head Up Games have been shortly discussed to illustrate our vision, and the experiences gained from creating and evaluating these games show that our original ambitions have been met to a large degree. In the next chapters we will go into more depth regarding the design, the creation and evaluation of Head Up Games.

REFERENCES

Bakker, S., Markopoulos, P., and Kort, Y. de. (2008). OPOS: an observation scheme for evaluating head-up play. In Proc. of NordiCHI '08 (pp. 33–42). Presented at the NordiCHI'08, ACM, Lund, Sweden.

Bekker, T., Sturm, J., Wesselink, R., Groenendaal, B., and Eggen, B. (2008). Interactive play objects and the effects of open-ended play on social interaction and fun. In Proc. of the 2008 Int. Conf. in Advances on Computer Entertainment Technology (pp. 389–392). ACM, Yokohama, Japan.

Benford, S., Anastasi, R., Flintham, M., Greenhalgh, C., Tandavanitj, N., Adams, M., and Row-Farr, J. (2003). Coping with uncertainty in a location-based game. Pervasive Computing, IEEE, 2(3), 34–41.

Benford, Steve, Crabtree, A., Flintham, M., Drozd, A., Anastasi, R., Paxton, M., Tandavanitj, N., Adams, M., and Row-Farr, J. (2006). Can you see me now? ACM Trans. Comput.-Hum. Interact., 13(1), 100–133.

Benford, Steve, Rowland, D., Flintham, M., Drozd, A., Hull, R., Reid, J., Morrison, J., and Facer, K. (2005). Life on the edge: supporting collaboration in location-based experiences. In Proc. of CHI '05 (pp. 721–730). ACM, Portland, Oregon, USA.

Brederode, B., Markopoulos, P., Gielen, M., Vermeeren, A., and Ridder, H. de. (2005). *pOwerball*: the design of a novel mixed-reality game for children with mixed abilities. In Proc. of the 2005 Conf. on Interaction Design and Children (pp. 32–39). ACM, Boulder, Colorado.

Brewster, S., Lumsden, J., Bell, M., Hall, M., and Tasker, S. (2003). Multimodal "eyes-free" interaction techniques for wearable devices. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 473–480). ACM, New York, NY, USA.

Bruegel, P. (1560). Children's Games. Retrieved February 23, 2009, from http://www.gamesmuseum.uwaterloo.ca/VirtualExhibits/Brueghel/aboutpaint.html

Cheok, A. D., Anuroop Sreekumar, Lei, C., and Thang, L. N. (2006). Capture the flag: mixed-reality social gaming with smart phones. Pervasive Computing, IEEE, 5(2), 62–69.

Clements, R. (2004). An investigation of the status of outdoor play. Contemporary Issues in Early Childhood, 5(1), 68–80.

Consalvo, M. (2007). Cheating: Gaining Advantage in Videogames (annotated edition.). The MIT Press.

Consolvo, S., Everitt, K., Smith, I., and Landay, J. A. (2006). Design requirements for technologies that encourage physical activity. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 457–466). ACM, Montréal, Québec, Canada.

Cummings, J. N., Butler, B., and Kraut, R. (2002). The quality of online social relationships. Commun. ACM, 45(7), 103–108.

Curtis, P. (2009, January 19). Internet generation leave parents behind. The Guardian. Retrieved from http://www.guardian.co.uk/media/2009/jan/19/internet-generation-parents

Fujiki, Y., Kazakos, K., Puri, C., Pavlidis, I., Starren, J., and Levine, J. (2007). NEAT-o-games: ubiquitous activity-based gaming. In CHI '07 Extended Abstracts on Human Factors in Computing Systems (pp. 2369–2374). ACM, San Jose, CA, USA.

Gray, P. The decline of play and the rise of psychopathology in children and adolescents. American Journal of Play 3, 4 (2011), 443–463.

Hellenga, K. (2002). Social space, the final frontier: Adolescents on the Internet. The changing adolescent experience: Societal trends and the transition to adulthood, 208–249.

Hendrix, K., Herk, R. van, Verhaegh, J., and Markopoulos, P. (2009). Increasing children's social competence through games, an exploratory study. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 182–185). ACM, Como, Italy.

Hendrix, K., Yang, G., Van de Mortel, D., Tijs, T., and Markopoulos, P. (2008). Designing a Head-Up Game for Children. In Proc. of the HCI'08 Conf. on People and Computers XXII (Vol. 1, pp. 45–53).

Lund, H. H., Klitbo, T., and Jessen, C. (2005). Playware technology for physically activating play. Artificial Life and Robotics, 9(4), 165–174.

Magerkurth, C., Cheok, A. D., Mandryk, R. L., and Nilsen, T. (2005). Pervasive games: bringing computer entertainment back to the real world. Comput. Entertain., 3(3), 4–4.

Magielse, R., and Markopoulos, P. (2009). *HeartBeat*: an outdoor pervasive game for children. In Proc. of the 27th Int. Conf. on Human Factors in Computing Systems (pp. 2181–2184). ACM, New York, NY, USA.

Manning, M.L. Play development from ages eight to twelve. In D.P. Fromberg and D. Bergen, eds., Play from Birth to Twelve: Contexts, Perspectives, and Meanings. Routledge, 2012.

Markopoulos, P., Read, J. C., Macfarlane, S., and Höysniemi, J. (2008). Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers. Morgan Kaufmann.

Marzano, S. (2008). Ambient Culture. In J. L. Encarnação and E. H. L. Aarts (Eds.), True Visions: The Emergence of Ambient Intelligence (1st ed. 2006. 2nd printing.). Springer.

Mueller, F., Stevens, G., Thorogood, A., O'Brien, S., and Wulf, V. (2007). Sports over a Distance. Personal and Ubiquitous Computing, 11(8), 633–645.

Nieuwdorp, E. (2007). The pervasive discourse: an analysis. Comput. Entertain., 5(2), 13.

Nova, N., Girardin, F., Molinari, G., and Dillenbourg, P. (2006). The Underwhelming Effects of Automatic Location-Awareness on Collaboration in a Pervasive Game. In Int. Confernce on Cooperative Systems Design (COOP 2006) (pp. 224–238).

Olson, C.K. Children's motivations for video game play in the context of normal development. Review of General Psychology 14, 2 (2010), 180–187.

Pielot, M., Poppinga, B., Heuten, W., and Boll, S. (2011). 6th senses for everyone!: the value of multimodal feedback in handheld navigation aids. In Proc. of the 13th Int. Conf. on Multimodal

Interfaces (pp. 65–72). ACM, New York, NY, USA.

Robinson, S., Eslambolchilar, P., and Jones, M. (2009). Sweep-Shake: finding digital resources in physical environments. In Proc. of the 11th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services (pp. 12:1–12:10). ACM, New York, NY, USA.

Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., Harris, E., Smith, H., Randell, C., Muller, H., O'Malley, C., Stanton, D., Thompson, M., and Weal, M. (2004). Ambient wood: designing new forms of digital augmentation for learning outdoors. In Proc. of IDC '04 (pp. 3–10). ACM, Maryland.

Salen, K., and Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals (illustrated edition.). The MIT Press.

Scarlett, W. G., Naudeau, S. C., Salonius-Pasternak, D., and Ponte, I. C. (2004). Children's Play (1st ed.). Sage Publications, Inc.

Seitinger, S., Sylvan, E., Zuckerman, O., Popovic, M., and Zuckerman, O. (2006). A new playground experience: going digital? In CHI '06 Extended Abstracts on Human Factors in Computing Systems (pp. 303–308). ACM, Montréal, Québec, Canada.

Short, J., Williams, E., and Christie, B. (1976). The Social Psychology of Telecommunications. John Wiley & Sons Ltd.

Soler-Adillon, J., and Parés, N. (2009). Interactive slide: an interactive playground to promote physical activity and socialization of children. In Proc. of the 27th Int. Conf. Extended Abstracts on Human Factors in Computing Systems (pp. 2407–2416). ACM, Boston, MA, USA.

Soute, I., and Markopoulos, P. (2007). Head Up Games: The Games of the Future Will Look More Like the Games of the Past. In Human-Computer Interaction – INTERACT 2007 (pp. 404–407).

Soute, I., and Markopoulos, P. (2009). Evaluating the effect of the tangibility of game elements on physical activity and social interaction in Head Up Games. In CHI '09 Extended Abstracts on Human Factors in Computing Systems.

Turkle, S. (2011). Alone Together: Why We Expect More from Technology and Less from Each Other (1st ed.). Basic Books.

Verhaegh, J., Soute, I., Kessels, A., and Markopoulos, P. (2006). On the design of *Camelot*, an outdoor game for children. In Proc. of the 2006 Conf. on Interaction Design and Children (pp. 9–16). ACM, Tampere, Finland.

Vygotsky, L. (1976). Play and its role in the mental development of the child. In J. S. Bruner, A. Jolly, and K. Sylva (Eds.), Play: Its Role in Development and Evolution. Basic Books, New York.

Chapter 3

Designing Head Up Games

This chapter is based on:

Soute, I., & Markopoulos, P. Designing for Embodiment in Outdoor Interactive Games for Children. In A. Nijholt (Ed.), Playful User Interfaces: Interfaces that Invite Social and Physical Interaction. Springer. To appear.

Soute, I., Lagerström, S., & Markopoulos, P. (2013). Rapid Prototyping of Outdoor Games for Children in an Iterative Design Process. In Proceedings of the 12th International Conference on Interaction Design and Children. New York: ACM. To appear.

ABSTRACT

This chapter focuses on the design process of Head Up Games. First insights gained during the design of several Head Up Games are presented. Next, we present a study in which we used high-fidelity prototypes in early stages of the design process. This is a deviation from the usual User Centered Design (UCD) or Game Design process, which typically start from low-fidelity prototypes, gaining in fidelity in an iterative process that involves user testing of the prototypes. We argue that, for the design of Head Up Games, it is more useful to start with developing high-fi prototypes, as the use of low-fi prototypes negatively influences the emerging game experience.

INTRODUCTION

This chapter discusses the methodology of designing Head Up Games. It is based on the experiences accumulated in designing a variety of games as part of this thesis and in a number of related student projects.

This chapter is divided into two parts: the first part discusses several methodological insights we gathered during the design of Head Up Games, which, by reporting them, can inform inexperienced game designers. These insights relate to several aspects of the design process, such as the idea generation phase; involving adult users instead of child users; and how to use children's tacit knowledge for generating new ideas for game rules.

The second part of the chapter examines in depth the impact of using prototypes in the design process; more specifically we argue that for Head Up Games design, high-fi prototypes should be employed. We discuss how this topic is approached in HCI and Game Design literature, and next focus on how formative user tests with paper or other low fidelity mock-ups are favored as a means to explore a design. In contrast, we argue that to properly design for and judge the added value of novel interaction styles in Head Up Games, designers need to create high-fi prototypes in early stages of the design process, so that users can truly experience the game play. This approach has been demonstrated in a study in which, in a rapid, iterative process four Head Up Games were designed, implemented and evaluated.

REFLECTIONS ON DESIGNING HEAD UP GAMES

The following section discusses and reflects on several methodological aspects. We argue that in the specific context of Head Up Games issues arise that have not been identified or reported earlier in related work. These issues relate to 1) idea generation, 2) getting to know your users 3) testing with adults and 4) tapping into children's tacit knowledge on games.

IDEA GENERATION

A game design process typically starts with an idea generation phase. Inspiration can come from anywhere, at anytime, and there are many methods and tools available for kick starting the designers' creativity. One widely known method is brainstorming, either alone or in a team. Though issues have been identified that can reduce the effectiveness of brainstorming (Stroebe et al., 2010), when prepared well, a brainstorm session can definitely assist in generating concepts; Rossiter and Lilien (1994) provide a set of general principles for conducting successful brainstorm sessions. More specifically for gaming, Fullerton et al. (2004, p. 142) give pointers for brainstorming (computer) games. In related work, many research projects in game design mention some form of brainstorming, e.g., (Kern et al., 2006; Valk et al., 2012).

Below we describe two brainstorm sessions that were held during the development of *F.A.R.M.* because they illustrate possible benefits and challenges when brainstorming for Head Up Games. For the setup of the brainstorms we followed the principles of Rossiter and Lilien (1994), which are: (a) brainstorming instructions are essential and should emphasize, paradoxically, number and not quality of ideas; (b) a specific and challenging target should be set for the number of ideas; (c) individuals, not groups, should generate the initial ideas; (d) groups should subsequently join and refine the ideas; (e) individuals should provide the final ratings to select the best ideas, which will increase commitment to the ideas selected; and, (f) the time required for successful brainstorming should be kept remarkably short.

F.A.R.M. | p. 69

First brainstorm session

In the first brainstorm session eight designers participated, with backgrounds ranging from industrial design to game design. The session was organized as follows: first, as the participants did not know each other, we played a few games to familiarize them with each other. As the participants were not familiar with the concept of Head Up Games, we introduced it to them. We asked the participants to individually think about the games they liked to play in their childhood. Next, the participants were divided in three groups and we asked them to discuss their childhood games and identify elements of these games that added to the appeal of the games. We then asked them to create a new outdoor game for children. The participants should at least provide details on how the game could be won, though more details, like specific game rules, how many players, where to play etc. were also encouraged. For inspiration we provided them with a set of commercially available board and card games; the participants could use these, or any other game that they knew themselves, as inspiration. At the end of the brainstorm session each group presented their final game concept and each participant picked out his or her favorite concept.

The results from this brainstorm were game concepts on paper; most of them included elements of tag, hide-and-seek, capture the flag or a combination of these. Furthermore, what we concluded from the game concepts generated in this brainstorm is that many games seemed to be fun already, even without technology. This insight prompted us to conduct a second brainstorm session, but change the setup; with the change we hoped to generate concepts that would more meaningfully include technology in the game.

Second brainstorm session

The setup of the second brainstorm session was similar to the first brainstorm session. However, instead of asking the participants to use childhood memories or existing games as inspiration, we gave them several possible technologies for outdoor games as inspiration. Based on earlier experiences designing Head Up Games we compiled a list of technologies that we deemed appropriate for outdoor use. They were: RFID, distance detection, accelerometer, and a rotation encoder. All participants of this workshop were industrial design students, who where familiar with these types of technology and also with participating in idea generation sessions.

Similar to the first brainstorm session, we started by giving the participants an individual task: all participants were seated around a round table, and in the middle of the table a set of papers was placed. Each paper was marked with one of the technologies. We asked the participants to randomly pick a paper from the table, quickly jot down a game idea on that paper, and put the paper back. These game concepts needed not be very elaborate or detailed. Next they could take a new paper and repeat the process. If the paper already contained an idea of one of the other participants, the participant could either start a new idea, or add to the idea already on the paper. After approximately half an hour many ideas had been generated this way.

Next, we grouped the participants in pairs of two. We provided them with two of the papers with ideas from the previous exercise and asked them to discuss and take inspiration from the strongest ideas to create a new, detailed concept of a game. This exercise took around ten minutes, next we regrouped the participants and provided them again with two of the papers of the previous session and repeated the process. Afterwards, the game concepts were presented, discussed and rated.

Reflecting on Idea Generation

As we had expected the brainstorm sessions rapidly generated many ideas and concepts. However, we observed that it can be difficult for participants to think "outside the box" with respect to meaningfully including technology in their concepts. Especially in the first session some of the ideas generated appeared fun enough by themselves, without needing to add technology. In the second session this seemed to have improved as technology was taken into account from the start. However, still some of the ideas suffered from a "technology push", in that the games would have been fun when taking the technology out and replacing it by a non-technical counterpart.

Furthermore, we observed that ideas generated in the brainstorms were very extensive with regard to the number of rules and details. From a game design perspective this is undesirable; from our experiences creating Head Up Games we know that games typically do not benefit from having many rules. However, from a brainstorm perspective it is a good outcome: apparently the context of the brainstorm sessions allowed the participants to continuously create and expand on concepts. We need to keep in mind that concepts are not games yet; they serve as inspiration for designers. In that process, the observation that participants too easily add rules to game concepts is important to acknowledge and take into account when further developing concepts into games.

INVOLVING CHILDREN IN EARLY DESIGN METHODS

The User-Centered Design (UCD) process focuses on the user's wants and needs for interactive technology. To gather the user's requirements in an early phase of the design process, many methods are available: for example, user surveys (questionnaires, interviews) or observations can be conducted. The role of the user in these methods is more or less passive, in that he or she only reacts on what the designer proposes. A more active participation of children in the design process is proposed by Druin (1999): the Cooperative Inquiry methodology is a set of techniques that put children in the role of co-researchers or co-designers. Scaife et al. (1997) put forward the notion of *Informant Design*. Though in this framework children are not seen as co-designers, they are acknowledged as valuable participants that contribute to the design process: children and adults can work together in design activities to generate input for the various stages of the design process. For eliciting children's requirements we have mostly involved children as informants; here we describe the methods we have applied in various projects designing Head Up Games.

Mission from Mars

Dindler et al. (2005) developed the method "Mission from Mars" to gather user requirements specifically for the design of children's technology. The method aims to

Figure 3.1 Setup for the session with the Martian. Left: the children's room with the video camera. Right: the Martian room with the audio equipment to distort the Martian voice

create a shared narrative space that allows researchers to get insights in the user requirements in an informal, fun setting for the children. First, the narrative is established: children are introduced to the story of a "Martian" who is eager to learn more about a specific subject; the Martian is ignorant on this subject because it is non-existing on Mars. As the Martian thinks that children on Earth are more knowledgeable on the subject, the Martian wishes to have contact to discuss and learn from them. Second, supported by the researchers, the children prepare for the encounter with the Martian. Finally, the children have the encounter with the Martian: a setup is installed, where the children can hear the Martian only, though the Martian can both see and hear the children, so that they can show what they have prepared. Practically, this means that the children are facing a video camera during the encounter; the children talk to a video camera and get feedback from the Martian through a set of speakers (see Figure 3.1). That signal is put through to the room of the Martian, where the researcher acting as the Martian can respond to the children. The voice of the Martian is distorted, to make him sound more "believeable".

The main reason for engaging in such an elaborate setup is to place children in the *expert* position, in which they feel free to share many details. The setting allows the researcher to ask 'stupid' questions about details that would have been impossible to ask in a conventional setting. For example, during the development of *Camelot* (Verhaegh et al., 2006), the Martian asked the children what a ball was, to which the children gave a serious and elaborate answer. In a post-hoc interview one of the children mentioned that in a conventional interview setting she would not have provided this level of detail, because she assumed adults to know what a ball is.

Camelot | p. 34

Dindler et al. (2005) used the method to gather insights for the creation of 'eBag', an electronic school bag. They applied the method with children 10-11 years. During the development of *Camelot* we applied Mission from Mars to obtain information on what games children prefer to play. We applied the method with children aged 7-9. Similar to Dindler et al. we concluded that indeed a significant amount of information is gathered using this method. Furthermore, Dindler et al. reflected on the credibility of the Martian narrative. Some of the children did not believe the story about the Martian to be true; however, this did not have an effect on the outcome of the study because the children played along anyway. In contrast, we observed that children from a younger age group did believe the narrative, and though the majority of the children enjoyed participating and communicating with the Martian, some of the children were quite anxious about meeting the Martian. So, we seem to have run into the limit of for what age group this method is appropriate: for younger children the method could arguably be too intimidating.

Collage making - a creative exercise

Another early user requirements gathering method is KidReporter (Bekker et al., 2003). In the KidReporter method, children are asked to undertake various activities that result in creating a newspaper with children's ideas on a certain topic. For example, children could take pictures and describe why they took these pictures and what is on them. Furthermore, children could interview each other, reporting on that, or independently write an article about a topic. The KidReporter method inspired us to do a similar activity during the Mission from Mars method: to inform the Martian about the games children liked, we asked the children to create a collage of their favorite games. As a preparing activity, we gave the children small cameras so they could first take pictures of their favorite games and subsequently use these pictures in their collages.

This idea worked out well: the children really made an effort to take photos of their favorite games; most children

Figure 3.2 Making the collage as preparation for the session with the Martian

documented the games they were playing that afternoon at school, and some children went as far as to stage all their favorite games after school hours so they could take photos of them. What we did not anticipate was that the act of making the collages, which we did the next day at school (see Figure 3.2), would generate a considerable amount of information. Each group was guided by an adult, and it turned out that, while the children were engaged in making the collages, they were very open to elaborate verbally on details of their favorite games. It was very easy for us to surreptitiously pose many questions to which the children answered freely. We attributed this to the fact that for the children the main activity was to make the collage, which they enjoyed, and they did not feel as though they were being interviewed.

During the development of a series of Head Up Games (see later in this chapter) we decided to again create collages with children to gather information. The main difference with the session described above was that this time we did not include the photo making activity - instead we brought crafting materials. Furthermore, the context was different: instead of children at school, we engaged with children at the scouting.

Based on our previous experience with making collages with children we expected it to be a good opportunity to simultaneously interview the children. However, totally unexpected, this time our experience was very different: in contrast to the school children, the scouts did not enjoy the activity, fooled around a bit and effectively did not provide us with any information.

Reflecting on this we argue that the context in which we executed the activity has a significant influence on the proceedings. At school, children are normally required to behave calmly and an activity such as making a collage is a welcome deviation from the normal school routine and thus perceived as a fun activity. In contrast, children go to a scout meeting to be playful and engage in games and play in the woods; it is a venue for the children to release pent up energy from a week's worth of attending school. In that context, an activity as collage making, which required the children to sit at a table and be relatively calm, was *not* seen as fun.

Furthermore, our experiences show that information can be generated at unexpected times: while preparing the Mission from Mars, we had not expected that making the collages would give as this much information; we merely saw it as a means for the children to prepare for the session with the Martian.

Observations

In most of the design processes for Head Up Games we have spent time observing children's free play in their natural context. In our experience this is a necessary activity, at least for inexperienced game designers. Though the methods described above will result in more and detailed information, they also take more time to prepare, execute and analyze; time that is not always available. However, we argue that it is important for a designer to familiarize him- or herself with the target audience, and observing children at play is a way to gather insight in the types of games they play, the language they use and the context in which the games are played. Not only will it give valuable insights for the design process, it will also help the designer/researcher to better prepare for evaluations with children of eventual prototypes of games.

Reflecting on involving children in early design methods

We have shortly highlighted three methods for gathering insight. What method best suits a design process depends on several factors: the amount of time available in the process, the desired level of involvement of the children, and, from a practical point of view, the accessibility to children. A method like Mission from Mars requires a substantial effort in time and resources to execute and we have seen that the method's suitability depends on the age of the children. Then again, if the aim of the process is to involve children as design partners or informants it is worth the effort to spend time with the children to build up a relationship for subsequent encounters.

Similarly, we argue that an activity as collage making can also be deployed as a requirements gathering tool, and arguably as a relationship catalyzer; though our experiences suggest that the context in which the activity is conducted must be carefully considered. Spending time with the scouts in a shared activity that better matches the scouts' context arguably would have been more informing for the design process.

Based on our experiences with Mission from Mars and the collage making activity, we argue that spending time with children in a fun, creative activity or a shared narrative can provide valuable insights for a designer. In general, it is advised to attempt to create a fun experience for children when involving them in a design process (e.g., Gielen (2008) Markopoulos et al. (2008)). We add to this observation that it is necessary to carefully select the right activity that matches the children's context.

Finally, we acknowledge the fact that given the timeframe of a design process it is not always possible to actively involve children, or alternatively, it is not possible to find a venue that allows for such active cooperation. For example, we found that it is not always easy to find for example a school willing to cooperate given the involvement we ask from them. Regardless the (desired) involvement of children, we feel that, especially for inexperienced Head Up Game designers, an effort should be made to at least (passively) observe children at play.

PLAYTESTING WITH ADULTS

Ideally a designer evaluates his prototypes or products with the intended end-users. However, from a practical point of view it is not always feasible to evaluate with children during *every* step of the design process. Testing with children takes time and effort to arrange, and to ensure that the children's time is well spent, it is important that most glitches in the games have been identified earlier. Therefore, during the development of many of the Head Up Games described in this thesis, we have asked adults to playtest intermediate prototypes. It is commonly acknowledged that we should not treat children as miniaturized adults, and as such evaluations with children merit from methods especially designed for children (Markopoulos et al., 2008). However, we would like to reverse that statement: can we treat adults as oversized children for the purpose of evaluating Head Up Game prototypes? Comparing our experiences of evaluating both with children as well as with adults we observed the following similarities and differences:

First, the behavior of adults before playing the game was different from children. Before the game children often behaved excited, eagerly anticipating the gameplay. In contrast, adults acted placidly, and some even assumed an air of "well, I'm only participating to do you a favor..." - adults were less excited about playing a game, or at least did not express this. Furthermore, we observed that adults patiently listened to our explanation of the game rules and game details, while at least half of the children did not bother to listen to the details once they had grasped the main goal of the game.

However, the moment a game started, instantly the behavior of adults changed and closely matched the behavior of children during game play: both groups became physically active, there was social interaction (shouting, cheering etc.) and adults responded similar to breakdowns in a game as children did. For example, in evaluations of *F.A.R.M.* we did not set a rule for the starting distance between the one player and the rest of the players. Adults responded in a similar fashion to this as children did; both commented that "it was unfair" if the distance was too close and both groups resolved the issue within seconds (see also next section).

F.A.R.M. | p. 69

After playing the game, when we asked players for feedback, there was again a noticeable difference between adults and children: adults were more fluent in providing feedback than children were, which is not really surprising. Children have not yet properly developed the ability to reflect on a meta-level and/or simply lack the vocabulary to do so (Markopoulos et al., 2008). Furthermore, there was a difference in the type of feedback given. Children mostly reflected on actual events of the playtest; though children generally did not have problems to detect and fix "broken" game rules, they did have trouble to give feedback on the game on a more abstract level. Adults did not have trouble doing this and also commented more on the tactics of a game; they readily provided many more rules that they thought would enhance the game play. We should note here that though adults will come up with many rules and they expect that these rules will enhance the game play, this is not always the case: in our experience simply adding more and/or more complex rules does not automatically improve the game experience of outdoor games. We argue that this is partially due to the fast pace of many of the games. There simply is less time to consider all these rules while playing the games. Also, children seemed to appreciate other challenges in games than adults: adults put more emphasize on developing play tactics in the game, and also favored rules that would support this. For example, after playtesting *F.A.R.M.* the adults suggested to add more tactics to the game. They proposed rules that for example would allow players to trick other players into losing their animals. In contrast, children seemed to be less concerned by this; the first indication for this is that at the start of the playtest, children were less inclined to listen to the game rules. Interestingly, this did not seem to have a big influence on the game play. Not understanding all the rules while playing a game inevitably did result in confusion for some of the children, but in general they would just continue, figuring out most rules while playing. This observation was reflected in the informal interviews: when asked what children favored most in the games, some children referred to the physical activity, other children mentioned the fact that they were playing this game with their best friend. So in contrast to the adults, children did not seem to have a need for more (tactical) challenges in the games.

Summarizing, we observed that it is indeed beneficial (and practical) to have adults playtest the games; the behavior adults display during play is similar to children's behavior in terms of physical activity, social interaction and reaction to the game devices. Observing adults playtest the game can therefore identify usability issues with the game devices (e.g., sounds being not clear enough) or issues concerning the rules (e.g., when situations occur in which the rules are inconclusive, or conflicting). However, we would certainly not advise testing with adults only; though the behavior of children and adults is similar, children experience and value games differently than adults, and this cannot be revealed by testing with adults only.

TAPPING INTO THE CHILDREN'S TACIT KNOWLEDGE ON WELL-PLAYED GAMES

Game Design literature (e.g., Fullerton et al. (2004), Salen and Zimmerman (2003)), states that it is impossible to design all the rules and mechanics of a game and predict the emerging game experience without playtesting. Therefore, Game Design literature stresses the importance of an iterative design process, in which designers playtest the games; based on the observations designers can improve the game play.

DeKoven (2002) argues that players instinctively know when a game is played "well". He states that a "well-played game" is impossible to define, as it is dependent on too many variables. However, the *experience* of a well-played game is familiar to every player. Hughes' (1983) research makes a similar observation: children intrinsically aim to play "nice", e.g., it is implicitly agreed that players will not hurt each other. Furthermore, Hughes, and also Salen and Zimmerman (2003, p. 130) suggest that some rules are implicit, ingrained by the social context in which children play the games. For example, a child playing a game like *F.A.R.M.* with his younger brother, would allow the younger child more leeway than he would were he playing with his best friend, who has roughly the same age.

We argue that a designer can make use of the observations stated above. First, we acknowledge the fact that a designer is not able to predict the game experience beforehand, and therefore is not able to design a definitive set of rules for a game in a single iteration of the design. Thus, we propose to purposefully design a limited, basic set of rules only. We expect that during playtesting situations will arise that will "break" the game, because the basic set of rules is insufficient. If such a situation arises, we propose to rely on the children's tacit knowledge of a well-played game and their ability to come up with a new or changed rule to fix the game play. If possible that rule will take effect immediately, which allows us to instantly reflect on the suitability of the rule.

During the design process of the Head Up Games we have encountered examples that this way of working is indeed useful for informing the design process. For example, while designing *F.A.R.M.* we did not explicitly state in the rules what the starting distance between the players should be. Upon start of the game, it immediately became clear to the players, that this distance had a big influence on the chances of winning for the player who was "it" (the *farmer*). Players commented on the unfairness of this situation and we discussed with the players how to improve this. The players

F.A.R.M. | p. 69

suggested to give the *farmer* some leeway; they argued that this was common in other games as well, and this largely solved the issue as we experienced immediately during the subsequent playtest.

Another example occurred during the evaluation of *Save the Safe* (Soute et al., 2009). We compared two types of game play: one with a digital (virtual) object and one with a physical object (a ball). Unexpectedly, the game with the ball ended very rapidly, because the first player grabbed the ball and headed to the Safe to end the game. Immediately, the opposing team started protesting that this was "unfair", since "you are not allowed to walk with the ball!". In fact, we had not imposed any rule stating such a thing, but many ball games indeed have such a rule: the player who has possession of the ball is forbidden to move. After a very short discussion - the winning team, at first reluctantly, agreed, since they saw too that there was no fun in playing a game that ended this abruptly - we agreed to impose the rule (not walking with the ball) for this game.

Save the Safe | p. 36

Reflecting on tapping into tacit knowledge

Concluding, we argue that we can make use of the observation that children are in fact domain-experts to our advantage for informing the design process. However, we should keep in mind that children are domain-experts regarding game play, though not regarding technology. It is mostly impossible for children to comprehend in what ways technology can be used in the game; and this can result in either children not being able to imagine including novel interaction styles in games, or alternatively, children imagining game interactions that are technically infeasible to implement. By having children create rules and immediately play them, we are certain that these rules are playable. Still, the "blue sky" suggestions of children, combined with observations of children playing the game, can provide valuable hints to a designer on what direction to take in the game design process.

CONCLUSION

In this part of the chapter we have raised and addressed several methodological issues on how to meaningfully involve children and also adults in the design process of Head Up Games. The discussion of these issues can inform other Head Up Game designers on how they could implement their design process.

USING PROTOTYPES

In the second part of this chapter we will focus on the role of prototyping in the design process of Head Up Games. The generally accepted approach in HCI and Game Design is to start with low-fidelity prototypes that through subsequent iterations gain in fidelity and start to resemble the intended product more closely. In the design process of early Head Up Games we adopted this approach. For example during the development of *Camelot*, we playtested three game concepts using simple paper cards and boxes to represent some of the game ideas. Similarly, during the development of *F.A.R.M.*, we playtested the game with adults using paper prototypes. Although these evaluations were successful at first sight in that they gathered a considerable amount of insight in the game play, the question arose whether or not the information gathered was valuable for informing the design process of Head Up Games.

Camelot | p. 34

F.A.R.M. | p. 69

In this section we first present related literature on the role of prototyping in HCI and Game Design. Next we present a study in which we follow a different approach: namely making use of high-fi prototypes from an early stage in the process.

THE ROLE OF PROTOTYPING IN HCI AND GAME DESIGN

HCI

A generally accepted design process in HCI is the User Centered Design (UCD) process. This process advocates the involvement of users in all stages of the design process to ensure that the end product is valuable in terms of usage and experience for the user. Typical for a UCD process is the fact that it is *iterative*, i.e. the product is iteratively created, tested, improved and refined. It generally starts with a user-requirements phase, in which users are interviewed or observed to gather requirements. Next, a first iteration, often a low-fi, e.g., a paper prototype, of the intended design is created, which is evaluated with users. Results of such an evaluation are fed back into the design process, the concept is improved, and the process of creating a new prototype followed by testing is repeated. Typically, each cycle sees an improvement of the fidelity of the prototype, meaning that each time it increases in resemblance to the intended end product.

A major theme in HCI methodology concerns prototype fidelity, i.e., how similar a prototype should be with respect to the intended interactive system (Virzi, 1989). The term 'prototype' is ambiguous at best: it can range from paper prototypes (paper-based mock-ups), or sketches, to functioning, interactive nearly-ready products. Prototypes can be employed in many stages of the design process in different roles; for designers to explore the design space, as a communication tool to convey ideas to other stakeholders and furthermore as an evaluation tool to test a concept with end-users (Avrahami and Hudson, 2002; Lim et al., 2008). As early as the eighties paper prototyping had been established as a way to efficiently explore the design space without incurring development costs, and to obtain early user feedback on selected aspects of the interaction design. Proponents of paper prototyping, (e.g., Rettig, 1994) argued that software prototypes are less suitable for iterative design for several reasons: for example, these prototypes take longer to make and developers may resist changes that are found necessary during user testing. Furthermore, feedback on working prototypes tends to also include feedback on details, e.g., fonts and colors that are less relevant during the early phases of design. Also, technical glitches can run a test session to a halt. Especially the idea that different feedback can be obtained during user tests depending on the fidelity of the prototype prompted numerous experimental works, see Sauer and Sonderegger (2009) for an overview.

Over the years, the concept of fidelity has been nuanced, and methodologists tend to distinguish between different aspects in which a prototype can be more faithful as a representation of an intended design. McCurdy et al. (2006) introduced the concepts of visual fidelity, depth and breadth of functionality, richness of interactivity, and richness of data model, which are orthogonal dimensions for characterizing a prototype. It is clear, that depending on the type of system designed and the concerns of a designer at any particular point in the design process, different types of prototypes will be needed and therefore different types of prototyping platforms.

Game Design

Literature on game design advises to adopt an iterative design process and to playtest often and early (Fullerton et al., 2004; Lundgren, 2008; Salen and Zimmerman, 2003). In this process, in contrast to the UCD process, game designers appear to be less concerned with involving their end users in all stages of the process; there is less emphasis on getting to know the user and gathering user's requirements. Instead, the concept generation phase is mostly attributed to the game designer, relying on his/her experience in this field. However, Game Design literature does emphasize the importance of iteratively designing the game, in combination with play-testing: as Salen and Zimmerman (2003) put it: "the act of play becomes the act of design". The general opinion is that the resulting play experience of a game cannot be predicted at the 'drawing'-table. A game designer designs the rules and mechanics of a game, but the resulting game experience is ultimately generated by a player playing that game (Costikyan, 2002). As such, to be able to properly judge the game experience, the game *must* be played. The design process proposed is to rapidly prototype a playable version of the game, starting with low-fi paper prototypes and increasing fidelity in subsequent iterations.

Game design literature concerns mostly traditional computer games leaning towards software tools, methods and development techniques that are common in software engineering, e.g., agile methods*, are put to use to develop computer games as well (Flanagan et al., 2005; Koivisto and Suomela, 2007). Manker and Arvola (2011) explore prototyping practices in computer game design and conclude that prototypes in game design are used in many ways: amongst others, to externalize specific parts of the game and as a 'shared representation' that serves as a communication tool. Ollila et al. (2008) present an overview of the use of prototypes in early pervasive game development, though the overview is limited to games on mobile phones. The authors present guidelines for selecting the right prototyping method, based on the type of project, the phase of the project, the type of game and the purpose of the project.

* http://en.wikipedia.org/wiki/Agile_software_development

DESIGNING OUTDOOR GAMES WITH AND FOR CHILDREN

A growing number of games for children presented in the research literature are designed to be played outdoors by groups and often involve physical and embodied interaction. Examples of these are *Ambient Wood* (Rogers et al., 2004), StarCatcher (Brynskov and Ludvigsen, 2006), *Camelot*, and *HeartBeat*. Related publications describe how these games were designed following a standard user centered design process in which children were involved at various points in an iterative process as informants and testers (Scaife et al., 1997). Specifically iterations with paper prototyping were made before gradually technology was introduced; a design process, thus, that is generally accepted both in HCI as well as Game Design.

Camelot | p. 34
HeartBeat | p. 37

LOW-FI PROTOTYPING VS HIGH-FI PROTOTYPING

As we became more experienced in designing Head Up Games, we started to question the validity of using low-fi prototypes in the process and we identified several issues.

First, by employing low-fidelity prototypes (e.g., paper prototypes, sketches) in a playtest the intended game mechanics are altered to such an extent that it is no longer possible to compare the play experience to playing a prototype of a game that does employ interactive technology. As a consequence, the feedback generated with the non-technical prototypes is reflecting on irrelevant game mechanics, which results in the design process optimizing towards a game that is playable as is, i.e. without interactive technology. Subsequently integrating technology degrades new interactive features to unconvincing post hoc add-ons that do not integrate well with the game. Furthermore, the development of interactive prototypes takes a considerable amount of time, leaving little time to evaluate the prototypes and no time at all to implement improvements and re-evaluate.

Notably in the cases above and in numerous other games where the design process is disclosed in related literature (e.g., *Ambient Wood* by Rogers et al. (2004), and *StarCatcher* by Brynskov and Ludvigsen (2006)), it is only reported in the last design iteration that a partially or fully working, playable prototype is created that covers a reasonable part of the game mechanics. Most commonly in this field, authors report only the final game design, how that was evaluated by users and suggested potential improvements of the 'final' game. None of the papers report or reflect on an iterative process for improving the game design using interactive prototypes. However, the relevance of user feedback and the relevance of the prototyping can best be appreciated by relating them to their impact further on in the design process. For this reason, when considering the value of prototyping for outdoor game design it is most appropriate to examine (1) several iterations of the design and (2) how the extent to which the prototyping medium and approach support (or hinder) the design process and (3) allow the designer to focus on designing the emerging play experience.

To better support the design process of outdoor, interactive games for children, we propose to skip the low-fi prototyping phase, and immediately create high-fi prototypes that are iteratively evaluated and improved. To support such a process we developed the RaPIDO platform (see Chapter 4). This platform was especially designed to support rapid prototyping of portable, multimodal interaction. It consists of several independently functioning, identical devices that offer a wide range of (mobile) interaction possibilities, for example, an RFID reader for detecting RFID tags; a sound processor and speaker for auditory feedback; RGB LEDs for visual feedback; and an XBee chip (radio) for wireless communication between devices. The devices have been created in such a way that they can function stand-alone (i.e. without wires attached) and are not dependent on an existing infrastructure such as a wireless network. Furthermore, the devices contain ample battery power to run at least 2-3 hours straight, while the size of the device remains wieldy for children's hands. A software library is provided to facilitate programming of the devices.

As stated earlier: prototyping is a broad term and can be interpreted in many ways. For the remainder of this chapter, unless explicitly otherwise stated, we ºdefine a high-fi prototype as a *working, interactive* prototype. Thus to frame it in McCurdy's dimensions (McCurdy et al., 2006) the breadth and depth of the functionality and the richness of interactivity of our prototypes are typically "high fidelity". We are less concerned with the level of visual refinement, i.e. the look and feel of the prototype.

STUDY SETUP

To satisfy our goal of using high-fi prototypes in a design process and exploring interactions as well, we decided to create three separate games. This way we could explore games that are quite different from each other and deploy a wider range of interaction styles.

We chose three game concepts that we thought suitable for implementation. One was a concept generated in an earlier creative workshop that had not yet been implemented. The second game concept was inspired by a game one of the authors played as a child. Finally, as the third game we selected to re-implement *Save the Safe*. All games were implemented on the RaPIDO platform (see Chapter 4) and took about three weeks to proceed from concepts to playable prototypes. The details of the games will be presented in the next section.

Save the Safe | p. 36

We organized three evaluative sessions, with children of a Scouting organization. The sessions were planned during the regular scout meetings on Friday evenings at 18:30h. As the sessions took place in November and December this meant that the games were played outside in the dark. The area that we playtested the games in was at the scout's home - in the woods. There was a grass area available, lighted by three huge lampposts; furthermore, the (unlit) woods surrounding the area were also available for play. It is important to note here that the children, as they were scouts, were used to playing outside in any weather condition. The children were aged 7 to 10 years old, mostly boys, and in total 16 children were part of the scout group - though the scout leaders informed us beforehand that not every child attended every meeting.

Before starting the sessions we informed the parents of the children of our plans and we obtained consent for their children participating in the sessions, and also for gathering video and photo material.

Each session was planned to take two hours - which was equal to the duration of the regular Scouting meeting. We planned to first let the children play the games, while we both observed directly as well as gathered video material for later reviewing. Furthermore, we planned to apply GroupSorter (see Chapter 5) to gather feedback from the

FOLLOW THE LIGHT

by Sussi Lagerström

At the start of Follow the Light all players line up on one side of the game field. The goal of the game is to be the first to reach the opposite side of the field. At the start of each turn a color, an animal and a numeric value are announced. If the particular color is present in the player's clothes, he/she is allowed to take a number of steps forward, the number of steps corresponds to the value that was announced. The size of the steps must be proportionate to the animal that was announced; e.g., if the animal is a mouse, only small steps can be taken. When the appropriate number and size of steps are taken, a turn is done and a next turn starts.

The devices announce the color, animal and step count: the LEDs indicate the color; the number of steps to be taken is indicated by the number of LEDs that switch on and the animal is represented by a sound. One of the researchers is at the end of playfield, also holding a device. She starts each turn by holding a tag to a device, which then broadcasts the color and step count information to the player devices. Subsequently, players must turn the rotation wheel, and based on the final orientation of the wheel, the device calculates which animal is selected and the audio starts playing. The player that first reaches the researcher wins the game.

children; this entailed that after playing all the games, we first asked the children individually to rank the games from most fun to least fun. Subsequently, we formed small groups (4-5 children) and repeated the ranking, though now as a group effort. By recording the ensuing discussion we aimed to gather qualitative insight in what elements of the games were experienced as fun (or not), and also identify areas for improvement. Furthermore, to elicit new ideas and inspiration for the further development of the games, we planned to let the children create collages of their favorite games. We expected that this would be a fun activity for the children, while it allowed us to pose some questions without putting the children in a real interview setting.

In short, the data we expected to gather was our own direct observation and experience of the game play; video data; ranking data and audio data from the group interviews. Because of the conditions of how and where the games were tested we did not expect to get high quality video footage, but we decided to capture it all the same as a backup in case we needed to review it later.

Our aim is to develop active, outdoor games for children, mainly focused on generating a fun experience. In this design process we decided not to quantitatively analyze the data, but rather take a "holistic" view on the data regarding the game play and emergent game experience, also taking into account the context in which the games are played. Based on the whole picture, we tried to identify areas of the games that could be improved that would enhance the emergent game experience. This could range from game interactions not being understood by players, to a game rule that resulted in unfitting interactions between players, or a technical issue with the prototypes.

Below are the accounts for each evaluation session. For each session we explain the games we designed and elaborate on our expectations of the game play before the session. Next we describe specific details for the session and we report the results of the evaluation, both in terms of the process as well as how and why we decided to change the games for the next iteration.

FIRST ITERATION

For the first session we created three games, which took about three weeks to design and develop. These games were implemented and were piloted with adults before evaluating them with children, to root out early usability issues. Note that for each game we defined a basic set of rules, and that a limited set of these rules was implemented on the devices. There are two reasons for doing this; first, we see the technology as a *supportive* element of the game, and,

F.A.R.M.

F.A.R.M. (Finding Animals while Running and Mooing) is an individual chasing and collecting game. At the start of the game, each player receives an assignment to collect a set of animals, e.g., a cow and two horses. The player that first completes his assignment wins the game. Players take turns in being the 'farmer'. At the start of a turn the farmer gets assigned an animal, which can be won by other players if they tag the farmer within 10 seconds. Players are allowed to trade animals to better match their assignment.

Players decide amongst themselves who starts the game as farmer and in which order the next players take up the role. The farmer grabs a token from a bag and holds it near his device for identification. Once the token is identified, the device starts making an animal sound and continues to do this for the duration of the tagging phase. When the sound stops and the farmer has not been tagged, he can keep the animal him/herself.

Figure 3.3 Impression of game play in first iteration. Above: at the start of a game. Below: chasing a player in F.A.R.M.

like in traditional games, many game rules are implicitly defined, which we argue do not need to be explicitly enforced by technology. To give an example: in a game of tag, players might before hand explicitly agree on the play area, but there is never a debate on *how* the actual tagging operation is defined; players implicitly know, from previous games of tag, or from other similar games. Second, from a design point of view, by defining only a basic set of rules we expect that there will be situations that are not properly covered by the rule set. When this happens, our strategy is to ask the children how they would prefer to change or add the game rules. Their response gives us valuable insight in how the game can be improved.

Games

An account of the three games that were developed, namely *Follow the Light*, *F.A.R.M.* and *Save the Safe* can be found on the pages indicated on the right. Instead of the dedicated hardware (the belts) that was developed for *Save the Safe*, now the game has been implemented using RaPIDO.

Follow the Light | p. 68
F.A.R.M. | p. 69
Save the Safe | p. 36

Expectations

Before we ran the first evaluation we had certain expectations regarding the game play, and also regarding the process.

First, we expected that the children would be able to understand the rules of every game. We also expected that there could be some minor technical problems and that situations would occur that would lead to discussion. In *Follow the Light* we expected debate concerning the sizes of the steps with regard to the animals.

In *F.A.R.M.* we expected debate about the starting distance between the farmer and the rest of the players. Further, we expected the players to form a strategy to win the game, i.e. in choosing to run or not to catch an animal.

In *Save the Safe* we expected that, since the key is invisible, there would be confusion about who has the key and that players would actually pretend to have it to lure opponents away from the actual player with the key.

Regarding the process, based on our earlier experience with working with children at schools and after-school settings, we expected that when making the collages the children would have an idea of what their favorite games are, they would discuss the games and different elements in the them, they would also discuss why they like a certain game. We expected the children to have fun while creating the collages and that it would give us insights into why children like certain (features of) games.

Procedure

Fourteen children (3 girls, 11 boys) were present at the scout meeting. As we had eight working devices, we split up the group to ensure each child got to play every game. The first group played *F.A.R.M.* and *Save the Safe*, while the other group were kept busy by the scout leaders. After playing the groups switched places (see Figure 3.3 for an impression of the game play). Next, because we were running out of time, the groups were merged and all together *Follow the Light* was played in teams of 2-3 children, sharing one device per team. After all games were played, we went inside where children individually filled in a ranking form, ranking all games from most fun to least fun. The children were divided in three groups and each group was asked to create a collage of their favorite game; we left it up to the children whether to reflect on the games we played or to create a collage of other games. Each group was led by a researcher or a scout leader who prompted the children while they were creating the collage to explain why a game, or a game feature was fun. We captured the process of making the collages on video. After the children went home we discussed the games with the scout leaders.

Results

From our observations we gathered that in general children enjoyed *F.A.R.M.*, though we also observed some issues. As we expected, immediately at the start of the game discussion started at what was an appropriate distance between the farmer and the rest of the players. The children easily agreed on what was a proper distance - they mentioned other games where this was done as well, so it was quite easy for them to discuss this. However, what we had not foreseen is that, though the children had an agreement on the distance, they would still very much try to cheat, which led to much argument. Furthermore, the sound indicating which animal was identified, only played on the device of the farmer. Sometimes the other children had difficulty hearing the sound. We did not observe any trading or other form of strategy making in the game. Players mostly tried to catch all the animals, including the ones that they did not need to collect. Finally, an animal could only be won by running very fast, and it was soon clear which of the children had an advantage over the others by being faster runners. A few of the slower children decided to not even bother running.

The children did not really seem to be enthusiastic about *Follow the Light*, some children even mentioned out loud that the game was boring. As expected, there was quite a lot of discussion on how big steps they were supposed to be. Because of lack of time the whole group had to play the game at the same time, which led to the game being played in teams of two to three players. Since the amount of players in a team got bigger, each team also had a larger variety of colored clothes on them; almost all the teams were allowed to take steps on every turn. The pace of the game was slow and there was hardly any excitement.

Save the Safe seemed to be appreciated most, though there were some complaints about it being difficult to transfer

INVADE THE CASTLE

by Sussi Lagerström

In Invade the Castle the players are divided in to teams of 2-3 players. The narrative of the game is based on invading a castle, and to do this teams have to collect three types of weapons (a catapult, an arrow and a shield) that are hidden in the woods. The first team to collect all three weapons wins the game.

The weapons are virtual entities: three scout leaders each get a device that represents a weapon, by repeatedly broadcasting a 'signature' signal. The scout leaders hide themselves in the woods before the game starts, and remain at their locations for the duration of the game. Next, each team gets a device, and at the base station, where one researcher is keeping track of the score, they select which weapon they are going to search for, by holding the corresponding tag to their device. Now, the device knows for which weapon to look. As children enter the woods, the device shows nothing, but once they come into range (approx. 30 meters) of the area where their selected weapon is located, the device turns red. As the team gets closer to the scout leader, the device gradually turns from red, through orange and yellow, to green. Once they are close enough (2 meters), they "acquire" the weapon by staying in range for 15 seconds, after which the device turns blue to confirm the acquisition. Next the team returns to the base station and they can select a new weapon to start searching for. Once a team has collected all weapons, the game ends.

the key to another player. Indeed, as expected, we observed players pretending to have the key. Players within the same team were trying to give the key to each other as to confuse the other team, although this did not work very well since the key did not transfer very easily. Right after the last game finished, the devices collectively crashed.

Based on the individual ranking data we calculated an overall ranking: (1) *Save the Safe*, (2) *F.A.R.M.* and (3) *Follow the Light*. We also calculated Kendall's *W*, a measure for the agreement between rankers. There was a high level of agreement between the rankings: $W = 0,75$. This preference was also supported by our own observations.

Doing the collages did not work with this group. The children had a hard time sitting down and concentrating on doing the collages. In the end they seemed to think that it was more fun to make a mess and playing around than discussing the games. The audio captured during the creation of the collages was unfortunately unusable (poor acoustics and several children speaking at once in a closed location).

Finally, we discussed with the scout leaders about what they usually do during the scout meetings, what kind of games they play. We learned that they usually play games invented by the scout leaders. The ideas for these games come from traditional outdoor games and strategic board games. The children are also used to playing out in the woods in the dark.

Implications for the design and process

After the evaluation session we briefly reviewed the video and audio data and we discussed our main observations. Based on that we decided what features of the games should be improved or replaced, taking into account whether that is achievable in the short development time available.

Based on the feedback of the children, our own observations and the rankings we decided to drop *Follow the Light* all together; we saw little possibilities to improve the game play. For *F.A.R.M.* we decided to rework the game for the next session. For *Save the Safe* we decided to only tweak some parameters to make the transferal of the key easier and fix the bug that caused the devices to crash. As we had decided to drop *Follow the Light*, we wanted to create a new game instead. Based on the remarks of the scout leaders about playing in the woods, we decided to create a game that could be played there.

With regard to the evaluation process, we decided to not again ask the children to create a collage. We discussed what happened with the scout leaders, and in retrospect, they thought it was not really surprising: the children go to the scout meetings in their free time on Friday evening and this is their venue to let go of all of the pent up energy of a whole week of sitting inside a class room. Thus, it was hard for them to sit and behave and take part in a creative exercise.

SECOND ITERATION

Games

First we describe for each game what changes were made to either game play and/or technology:

Invade the Castle. This game was developed to replace

Follow the Light. The idea for *Invade the Castle* came partly from discussions with the scout leaders (playing out in the woods in the dark, searching for and collecting things), some game features were inspired by *Camelot* (Verhaegh et al., 2006), and the idea of how to search came from the idea of using avalanche searching techniques (player gets to know when he/she gets closer to the target). Since both *F.A.R.M.* and *Save the Safe* were games where children have to run quite a lot we also decided to design a game where less running was needed and the children had to concentrate more on problem solving.

F.A.R.M. Though we acknowledge that a bit of cheating is often part of a game, and can even add to a positive game experience, we observed in the previous session that children cheated many times with respect to the distance that should be observed at the start of each turn; and children were getting annoyed because of this. Therefore we decided to program the devices that they would enforce a minimal distance between the players, before starting a turn: as long as the other players were too close to the farmer, all devices displayed a red color. Once the distance was right, the devices turned green and the game could advance. Furthermore, as we had observed that not all players heard the animal sound, we changed the game such that once an animal was selected, the corresponding sound was played on all devices.

Save the Safe. We fixed the bug that caused the devices to crash.

Expectations

For this session we were expecting that the changes we made to the games would make them easier to play for the players. In *F.A.R.M.* we expected that the children would not try to cheat as much. We also expected that finding the right distance between the farmer and the chasing players would be easy for the children. In *Save the Safe* we expected that the game flow would be like last time, since no changes were made. For *Invade the Castle* we were not sure if the players would understand how to find their targets, if they would have some problems while doing so or if they would find them at all.

Figure 3.4 Playtesting F.A.R.M. in the snow

Procedure

This time ten children (2 girls, 8 boys) were attending the scout meeting. The weather conditions were different than in the previous session: it had snowed and it was freezing (see Figure 3.4). Again we split up the group in two: the first group played *F.A.R.M.* and *Save the Safe*, while the second group was keeping themselves busy in a snowball fight. After the first group finished playing the games, the groups were switched. After both groups had played the two games, we went inside for a short break to warm up. Next, we played *Invade the Castle* with the entire group.

F.A.R.M. and *Save the Safe* both were played in an open playing field and we were able to video tape the sessions as well as observe directly ourselves. *Invade the Castle* was played in the woods, which made it nearly impossible to capture on video (see Figure 3.5 for an impression of the researcher's view from the base station). One of the researchers remained at the base station to track the progress of the game; the other researcher went into the woods with the children and walked around observing as much of the game play as possible.

When *Invade the Castle* was finished, we asked the children to rank the games from most fun to least fun. Again the group was split up in two and a researcher shortly interviewed each group. Afterward, the children went home and we reviewed the games with the scout leaders.

Results

In *F.A.R.M.* the players had difficulties using the tags and in some cases they needed our help. We realized that

Figure 3.5 Poor visibility while playtesting Invade the Castle

the RFID reader is placed exactly below the spot where a child's thumb is located when holding the device, which makes it difficult for a child to place the tag in the right place. Because of the cold weather children were wearing gloves, which made it even harder to handle the tags. The game flow was interrupted every time a player struggled to get the RFID tag identified. Finding the right distance between the farmer and the rest of the players was easier this time.

Though we fixed the bug that caused the devices to crash, we inadvertently did not change the parameters for the transferal of the key. Still, the players seemed to enjoy playing *Save the Safe*.

Some teams had a hard time finding the weapons in *Invade the Castle*, though at the end of the game each team managed to obtain at least one weapon. What we observed was that the children would walk in range of the weapon, and that their device started slowly blinking red; what happened was that they were really on the edge of the range, and the devices were quite sensitively programmed for this, immediately switching off the LEDs once they were out of range. This confused some of the children who did not understand which way to walk and thought their device was not working properly. Once we demonstrated how to handle the situation, the children had less trouble finding the weapons. Furthermore, the LEDs of the devices of the scout leaders were switched on, resulting in children only partly relying on their own devices to hunt, but as soon as they spotted the lights of the leader's devices they would go in a straight line towards the scout leader.

With regard to the ranking of the games in this session, the children were much less in agreement. The combined rank resulted in (1) *Invade the Castle*, (2) *Save the Safe* and (3) *F.A.R.M.*. However, there was little agreement (Kendall's $W = 0{,}31$). On closer inspection of the data we concluded that *F.A.R.M.* was ranked lowest, but there was tie for *Save the Safe* and *Invade the Castle* for first place.

In the interviews children told us that all of the games were fun; it was by comparison only that *F.A.R.M.* was the least fun. Further, the children appeared to have some trouble to reflect on their experiences and to discuss why one game was more fun than the other: *"I simply liked that game. Just because."*. Though one child could express his preference for *Save the Safe* with regard to the game play: "In *F.A.R.M.* it was just a lot of chasing. And in *Save the Safe* it was chasing too, but at the same time you also had to search too (for the key)".

Implications for the design and process

Because handling the tags in our opinion interrupted the game's pace in *F.A.R.M.* we decided to redesign that part.

As we made the mistake of not tweaking the parameters of *Save the Safe*, this remained our goal for the next session.

As we were programming the devices for *Invade the Castle* for this session, we had not expected that the LEDs being on would have such an effect on the game play, probably because we did our own development and testing during the day, and of course in the dark a light is much more showy than during daylight. Thus we decided to switch it off. Furthermore, we decided to improve the interaction mechanism for providing feedback when entering the range of the weapons.

Interviewing the children informally at the spot worked fine for us, though the question rises how valuable the information retrieved from the interviews is. Mostly, the interview data supported our own observations of the game play, there were only very few cases where new insights were obtained. Still, from a research point it is preferable to have data that supports our observations, to be able to triangulate the findings.

THIRD ITERATION

Games

F.A.R.M. was adapted so that tags were not needed to start a turn. Instead, when the players were properly aligned, the device would automatically select an animal.

Save the Safe. *Save the Safe* was changed for easier transfer of the key.

Invade the Castle. The devices of the scout leaders no longer showed a light once the game has started. The range for which the tracking devices pick up a signal was extended, and the interaction for showing the range was made more robust.

Expectations

In *F.A.R.M.* we expected that the game pace would be higher since they would now be able to play the game without our help.

In *Save the Safe* we expected that the flow of the game would change since the key transferred better now.

In *Invade the Castle* we expected that there would be less confusion for the players during tracking down the weapons.

Procedure

14 children (2 girls, 12 boys) attended this scout meeting. Again, we split the group up, though this time into three subgroups. Each group played *Save the Safe* and *F.A.R.M.*, while the other groups remained indoors with the scout leaders. We did not randomly divide the groups, but made a split based on age, to see if this would affect the game play in *F.A.R.M.*. This time it was raining quite heavily, though for the scouts the weather did not seem to be a problem; they normally also play outside when it rains. After all groups had played both games, we went indoor for a short break. After the break we played *Invade the Castle* with all children. Like in the previous session one of us was in the forest with the children observing the game flow, while the other remained at the base station. In contrast to the previous session, when the researcher at the base station could see some of the children playing, this time the woods were pitch dark and from the base station the game play could not be observed.

After the game play we asked the children to rate the three games from most fun to least fun. Because of time running out we managed to interview ten out of 14 children. Then we thanked the children for the participation in our evaluations and awarded them with a 'diploma' as game designers. Finally, we again discussed the games with the scout leaders after the children went home.

Results

As in the previous session children who enjoyed running and chasing liked playing *F.A.R.M.*; it was less popular for the other children, though the difference seemed less than in the previous session, probably because younger children were not asked to compete with the older children. With the removal of the RFID tags the game play went a lot smoother and the children were able to play the game without needing our help.

Save the Safe remained popular, though now the settings for transferring the key seemed a bit too easy.

Invade the Castle was again popular. This time it was very dark in the woods, and now the lights of the scout leaders' devices were switched off, the children really needed to navigate using their own device to collect the weapons. As the devices now responded more reliably, the children were able to do this. One of the scout leaders mentioned that he was hidden underneath a bush, and he heard the children walk in circles around him, before identifying his exact location. We also observed that two groups found their targets rather easily, whereas the other two had some problems. Again, after briefly demonstrating the workings of the devices they understood it better and they could find their target.

The ranking data resulted in the following combined ranking: (1) *Invade the Castle*, (2) *Save the Safe* and (3) *F.A.R.M.*. This is identical to the ranking of the previous session, though this time the children were more in agreement: Kendall's $W = 0{,}53$.

Implications for the game design

Though this was the last evaluation we reflect here on how we would improve the games based on the evaluation.

F.A.R.M. seems to have reached its limits. In its current form it is a game that appeals most to children that enjoy intense physical activity, but there is no challenge for the children who do not. In the first session, when the game was new, these children were more inclined to play along, but once it became clear that the faster children had a distinct advantage, their participation decreased. One interesting opportunity we see for this game, is to implement some sort of *skill balancing*, i.e. making it harder to win for the faster children and at the same time making it easier to win for the slower children. That way the challenge is balanced more equally which might make the game appealing for a broader range of children.

Though the settings for transferal of the key in *Save the Safe* still need to be further optimized, we think that the game is well balanced. Similar to *F.A.R.M.* this game requires physical activity too, but now as a group effort so there is less need for an individual player to run all the time. Also, the invisible key adds to the experience of the game.

Finally, the main feature of *Invade the Castle*, searching using the devices, clearly appealed to the children.

REFLECTIONS

The experience of creating and evaluating these games have generated insights on different levels, namely (1) on the rapid, iterative, design process (2) on evaluating with children in this particular setup and (3) on what interaction mechanisms and technology are appropriate for interactive, outdoor games for children.

Design process

Most design changes that we implemented based on our direct observations of the game play and comments of the children, concerned directly the play and interaction functionality. Seemingly small details, like the duration of some interactions, influenced heavily the emerging game experience, showing the inadequacy of evaluating a mock-up of the game interactions for example when playtesting functionality with Wizard of Oz interventions. For example, the virtual key in *Save the Safe* could have been prototyped using a piece of paper or other small physical object. However, it is easy to see that this would alter the game: a physical object is clearly visible to the other players, especially when passing it around between players, so the element of guessing which player actually possesses the key (as is present when the key is virtually represented) is taken out of the game completely. And particularly that feature of the game turned out to be the most fun part. Thus, we argue that instead of playtesting with paper prototypes, it is best to immediately focus on the actual, working, interactions.

With regard to the time it normally costs to develop a working prototype, we conclude that with the RaPIDO platform we were able to bring this time drastically down. Over the course of six weeks we were able to develop, evaluate and improve four games. Furthermore, because it was relatively fast and easy to create working prototypes, the platform allowed us to play around with the technology, and thus freely explore the design space.

Because of the high pace of iterations - we took one to two weeks to develop new iterations of the games - we did not have much time to run an in depth analyses of the results of the evaluations (e.g., run a structured observation, or content analysis of the interviews). Initially we thought we would have time to analyze the interview data, after the first session it was immediately clear that we would not. Further, the interviews did not yield as much information for improving the games as we had expected beforehand; directly observing the game play was much more effective. Nonetheless, the little information that was deduced from the interviews was useful for triangulating our findings from the direct observations.

Another benefit of rapidly iterating over small changes is that it becomes easier to observe the impact of a small feature change. We argue that this way the design process becomes a “self-steering” process: if based on an observation a wrong conclusion was drawn and subsequently a wrong decision regarding the game mechanics is implemented, the next session will immediately show the (negative) effect and the design decision can be undone quickly.

Testing early and often in the design process makes sure that as a designer you do not “fall in love” with your own (features of the) games. After only one week of implementing a game, it is much easier to toss a feature in favor of an improved version or abandon a game altogether. In contrast, if one has taken months to implement a game, it is much more difficult to part from it, if at an eventual user test it turns out that certain features do not work out as expected.

Iterative testing with children

We experienced that repeated evaluations with the same group of children has a few distinct advantages: first, we got to know the children, which made it easier to interpret observations. For example, a child that behaves in a certain way may or may not do that as a consequence of playing the game and it is relatively hard to tell the difference from a single observation only. However, when observing the same children over time, as a (game) designer it becomes easier to tell which behaviors can be attributed to a child and which might be the result of playing a certain game.

The second advantage is that the children got to know us and because of that gained confidence and were at ease in their interactions with us. An often argued side effect of the children getting acquainted to researchers is that the power imbalance, that might exist between a child and an adult (Hennessy and Heary, 2005; Markopoulos et al., 2008), is lessened. In fact, we even experienced this to the extreme; the power imbalance was *reversed* in a sense that we had to try hard to assert ourselves on the participants, simply to get and hold their attention. We attribute this to two causes: first, as we indicated, the children became familiar with us; and second, because the groups that we handled are quite big and the children clearly outnumbered the researchers, children did not feel at all intimidated by the two researchers (in contrast to evaluations where only one or two children are present). In the field of interaction design and children it is commonly advised to take measures to redress the power balance between adult researchers/designers and children. In our case, we eventually had to claim a leading role in order to quiet the group down, and make sure they were all paying attention. This did not seem to have a negative impact, possibly because the children equated us to their scout leaders and they too addressed the children in this manner.

This brings us to another observation: when observing "in the wild" it is important to adjust to the context of an evaluation (see also Rogers (2011)), and more specifically how an evaluator should interact within that context. In our case this meant that we positioned ourselves in the roles of scout leaders. Related to our observation above is the notion that one evaluation method cannot simply be transplanted from one context to another. Certain patterns of behavior have been established between the children and the scout leaders and as a researcher we argue that you should be aware of this and plan your evaluation accordingly. An example of this is the observation that using collages to elicit information from children as a method did not work well in this particular context, simply because the children were not used to sedentary activities within this context. In contrast: we have applied the same method earlier in a school context where it worked well.

The age of the children ranged from 7 to 10 year old. This is something we had not anticipated, but in the end had to adjust to: for some games it might occur that the challenge for a seven year old to compete with a ten year old becomes too high, resulting in a negative game experience for the seven year old, and maybe even for the 10 year old, as the competition is too low for him. We observed this during the playtest of one of the games, and later adapted for this by not randomly mixing the children but instead sorting them by age group. Then the chances for winning the game became more equal for all players, resulting in a better game experience.

Interaction in outdoor games

In total we designed, implemented and evaluated four different games. In these games we used a variety of interaction styles and technologies, but the one technology most commonly used was the radio communication. We used it for two purposes: for communicating game events between devices, so they could appropriately respond to what was happening in the game with respect to other players. Further, we used the radio technology for getting a rough estimation for distances between devices (and thus players). Both features contributed significantly to the novelty of game play, as it allowed us to introduce features in games that have no similarity to features in traditional outdoor games. An example is the virtual key in *Save the Safe*, which was transferred between players based on proximity.

Furthermore, for feedback to the players we often used auditory, visual and tactile cues. We found all modalities appropriate for supporting outdoor games, though that does not automatically imply that every style of using it is appropriate in the context of outdoor gaming. To give an example: in *Follow the Light* we needed to convey to the players how many steps they could take in a turn. At first we implemented it by letting the LEDs blink, each blink accounted for one step. However, this enforced the players to be paying attention to their device at a specific frame of time within the game. Also, the information is volatile, once it is shown it is gone. So, a moment of distraction, for example when talking to a team player (which we want our games also to encourage!) would result in the loss of game information. Therefore, we redesigned that part of the game to have the LEDs continously shine; the number of LEDs switched on corresponded to the number of steps. This was a more persistent way of showing the same information.

Another technology we made heavy use of was RFID technology. Each device is equipped with an RFID reader, which allowed us to program the devices to detect objects tagged with an RFID tag. Though we used it moderately in the games themselves, we employed the RFID tags mostly for setting up the games.

In contrast to other research prototypes, that often have a limited lifespan and/or are quite sensitive with respect to their environment, we were quite satisfied with how the devices performed: they were quite reliable, even in harsh conditions (snow, rain, low temperature, and rough handling).

CONCLUSION

We described in depth a design/evaluation process that deviates from the generally accepted way of using prototypes in HCI. Instead of gradually increasing the fidelity of the prototypes, we went ahead and immediately created high-fi (with respect to interactivity) prototypes. We argue that, for games involving physical activity, outdoor play, groups of players and embodied interaction, it is virtually impossible to test with paper prototypes as the lack of interactivity distorts the game dynamics intended by the designer and leads to very different play experiences. Moreover, we argue that, in the case where children are involved, it becomes more apparent, as children might be less able to reflect on the impact of interactivity and the resulting the game dynamics without actually experiencing it.

There is little research in the IDC field regarding the effects of media on the design process. For example, Hanna et al. (2004) found that children aged 8-9 can provide useful feedback to low fidelity representations of computer game concepts (specifically they used text accompanied by five screen shots). Still, that study concerned and seems most applicable to traditional video games where a script of the storyline and some scenes are sufficient to provide a description. It seems less applicable to cases where rather than the story it is the playful and embodied interactions between players that make up the game. Another rare example, where the impact of early design representations are considered by Kindborg (2002) compared the feedback children aged 10-11 can give for low fidelity and high fidelity representations of programming logic; he argues that children are better able to provide narrative accounts of software logic with non interactive representations, and were better able to consider alternative behaviors with the computational prototype. In all, it appears that the importance of considering how different prototyping media can or cannot support interaction design for and with children has not been explored sufficiently in this field.

Our work focused on the design of active, outdoor games, but we believe that similar arguments extend to tangible and embodied interaction in general. Future research should explore how design processes outside the game domain are or are not served by early use of such high-fi, rapid prototyping technologies like RaPIDO.

Based on our experiences we conclude that the process we followed is very suitable for games, where the emerging game experience is not only a result of interactive technology, but also of the context the game is played in, and other existing game rules; these games need to be really experienced and cannot be tested with lower fidelity prototypes. Arguably, a similar process might be valuable for other interactive systems that are to novel to users and are designed to change behaviors in users. To valuably generate feedback from users on such systems, they have to be tested and experienced “in the wild”.

RECOMMENDATIONS FOR DESIGNING HEAD UP GAMES

In this chapter we have focused on the design process of

Head Up Games. Based on our experiences we can now present several recommendations that can inspire and inform first-time Head Up Game designers.

The first recommendation is the most radical, as it deviates from the generally accepted way of involving low-fi to high-fi prototypes in the design process. Instead, we emphasize the necessity of using high-fi prototypes from an early stage of the design process: these games really need to be played with working technology to asses the effect of the game design and technology on the game experience.

Furthermore, we suggest starting the design process with designing and implementing a limited set of game rules and rely on the players' innate ability to detect and fix a broken game.

We also discussed the process of engaging children in creative activities to gather insight in children's requirements for games. The amount of information gathered is dependent on the effort put in the activities. The Mission from Mars method is very time consuming, but gathers a considerable amount of feedback. However, the method relies on a narrative that might not be suitable for all children. We furthermore discussed making collages with children. This method was a success in a school context, leading us to repeat it during evaluations with scouts. Unexpectedly, in that context the method failed, suggesting that the context of the evaluation plays a large role in its success rate. Finally, we suggest that at least designers should make an effort to observe children at play, if time is too limited to execute the methods described above for gathering children's requirements.

Finally, we recommend to playtest with children as often as possible. Our experience suggests that some issues can also be identified by adults. Indeed, we would recommend to let adults playtest intermediate prototypes of the games to root out early usability issues.

In short:

- Use high-fi prototypes from an early stage in the design process. Head Up Games really need to be played with working prototypes, most notably with regard to the game interaction, to assess the effect of the game design and technology on the emergent game experience.

- Start with an incomplete set of game rules. Rely on the players' innate ability to detect and fix a broken game to fill the gaps.
- Engage with children in a fun, creative activity as a way to facilitate discussion. Though be aware that 'fun' is context dependent.
- Adopt an iterative process and playtest with children as often as possible. To prepare for these sessions, and/or to test intermediate designs, playtest with adults too.

REFERENCES

Avrahami, D., and Hudson, S. E. (2002). Forming interactivity: a tool for rapid prototyping of physical interactive products. In Proc. of the 4th Conf. on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (pp. 141–146). ACM, New York, NY, USA.

Bekker, M., Beusmans, J., Keyson, D., and Lloyd, P. (2003). KidReporter: a user requirements gathering technique for designing with children. Interacting with Computers, 15(2), 187–202.

Brynskov, M., and Ludvigsen, M. (2006). Mock games: a new genre of pervasive play. In Proc. of the 6th Conf. on Designing Interactive Systems (pp. 169–178). ACM, University Park, PA, USA.

Costikyan, G. (2002). I have no words & I must design: Toward a critical vocabulary for games. In Computer Games and Digital Cultures Conf. Proc. (pp. 9–33).

DeKoven, B. (2002). The Well-Played Game. iUniverse.

Dindler, C., Eriksson, E., Iversen, O. S., Lykke-Olesen, A., and Ludvigsen, M. (2005). Mission from Mars: a method for exploring user requirements for children in a narrative space. In Proc. of the 2005 Conf. on Interaction Design and Children, pp.40–47.

Druin, A. (1999). Cooperative inquiry: developing new technologies for children with children. In Proc. of the SIGCHI Conf.on Human Factors in Computing Systems: The CHI Is the Limit (pp. 592–599). ACM, Pittsburgh, Pennsylvania, United States.

Flanagan, M., Howe, D. C., and Nissenbaum, H. (2005). Values at play: design tradeoffs in socially-oriented game design. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 751–760). ACM, New York, NY, USA.

Fullerton, T., Swain, C., and Hoffman, S. (2004). Game design workshop. Focal Press.

Gielen, M. A. (2008). Exploring the child's mind – contextmapping research with children. Digital Creativity, 19(3), 174–184.

Hanna, L., Neapolitan, D., and Risden, K. (2004). Evaluating computer game concepts with children. In Proc. of the 2004 Conf. on Interaction Design and Children: Building a Community (pp. 49–56). ACM, Maryland.

Hennessy, E., and Heary, C. (2005). Exploring children's views through focus groups. In Researching Children's Experience: Approaches and Methods (pp. 236–252). Sage Publications Ltd.

Hughes, L. (1983). Beyond the rules of the game: why are Rooie Rules nice? In F. E. Manning (Ed.), The World of Play. Leisure Press.

Kern, D., Stringer, M., Fitzpatrick, G., and Schmidt, A. (2006). Curball - a prototype tangible game for inter-generational play. (S. M. Reddy, Ed.). Ieee Computer Soc, Los Alamitos.

Kindborg, M. (2002). Comics, Programming, Children and Narratives. Interaction Design and Children, s. 93–109.

Koivisto, E. M. I., and Suomela, R. (2007). Using prototypes in early pervasive game development. In Proc. of the 2007 ACM SIGGRAPH Symposium on Video Games (pp. 149–156).

Lim, Y.-K., Stolterman, E., and Tenenberg, J. (2008). The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. ACM Trans. Comput.-Hum. Interact., 15(2), 7:1–7:27.

Lundgren, S. (2008). COVER STORY: Designing games: why and how. interactions, 15(6), 6–12.

Magielse, R., and Markopoulos, P. (2009). *HeartBeat*: an outdoor pervasive game for children. In Proc. of the 27th Int. Conf. on Human Factors in Computing Systems (pp. 2181–2184). ACM, New York, NY, USA.

Manker, J., and Arvola, M. (2011). Prototyping in game design: externalization and internalization of game ideas. In Proc. of the 25th BCS Conf. on Human-Computer Interaction (pp. 279–288). British Computer Society, Swinton, UK.

Markopoulos, P., Read, J. C., Macfarlane, S., and Höysniemi, J. (2008). Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers. Morgan Kaufmann.

McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B., and Vera, A. (2006). Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 1233–1242).

Ollila, E. M. I., Suomela, R., and Holopainen, J. (2008). Using prototypes in early pervasive game development. Comput. Entertain., 6, 17:1–17:17.

Rettig, M. (1994). Prototyping for tiny fingers. Commun. ACM, 37(4), 21–27.

Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., Harris, E., Smith, H., Randell, C., Muller, H., O'Malley, C., Stanton, D., Thompson, M., and Weal, M. (2004). Ambient wood: designing new forms of digital augmentation for learning outdoors. In Proc. of IDC '04 (pp. 3–10). ACM, Maryland.

Rogers, Yvonne. (2011). Interaction design gone wild: striving for wild theory. interactions, 18(4), 58–62.

Rossiter, J. R., and Lilien, G. L. (1994). New "brainstorming" principles. Australian Journal of Management, 19(1), 61–72.

Salen, K., and Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals (illustrated edition.). The MIT Press.

Sauer, J., and Sonderegger, A. (2009). The influence of prototype fidelity and aesthetics of design in usability tests: Effects on user behaviour, subjective evaluation and emotion. Applied ergonomics, 40(4), 670–677.

Scaife, M., Rogers, Y., Aldrich, F., and Davies, M. (1997). Designing for or designing with? Informant design for interactive learning environments. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 343–350). ACM, Atlanta, Georgia, United States.

Soute, I., Bakker, S., Magielse, R., and Markopoulos, P. (2013). Evaluating player experience for children's outdoor pervasive games. Entertainment Computing, 4(1), 25–38.

Soute, I., Kaptein, M., and Markopoulos, P. (2009). Evaluating outdoor play for children: virtual vs. tangible game objects in pervasive games. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 250–253). ACM, Como, Italy.

Stroebe, W., Nijstad, B. A., and Rietzschel, E. F. (2010). Chapter Four - Beyond Productivity Loss in Brainstorming Groups: The Evolution of a Question. In Mark P. Zanna and James M. Olson (Ed.), Advances in Experimental Social Psychology (Vol. Volume 43, pp. 157–203). Academic Press.

Valk, L. de, Rijnbout, P., Bekker, T., Eggen, B., Graaf, M. de, and Schouten, B. (2012). Designing for playful experiences in open-ended intelligent play environments. In Proc. of IADIS Int. Conf. Game and Entertainment Technologies (pp. 3–10).

Verhaegh, J., Soute, I., Kessels, A., and Markopoulos, P. (2006). On the design of *Camelot*, an outdoor game for children. In Proc. of the 2006 Conf. on Interaction Design and Children (pp. 9–16). ACM, Tampere, Finland.

Virzi, R. A. (1989). What can you learn from a low-fidelity prototype? In Proc. of the Human Factors and Ergonomics Society Annual Meeting (Vol. 33, pp. 224–228).

Chapter 4

Creating Head Up Games with RaPIDO a platform for rapid prototyping of mobile outdoor games for children

This chapter is based on:

Soute, I., & Markopoulous, P. Creating Head Up Games with RaPIDO: a platform for rapid prototyping of mobile outdoor games for children. Submitted to ACM Transactions on Computer-Human Interaction (TOCHI)

ABSTRACT

Outdoor multi-player games involving proximal social interaction and physical activity are an emerging class of interactive applications attracting growing research interest. We argue that prototyping tools to support designers in this domain do not address the specific domain requirements making prototyping interaction and game design concepts both ineffective and inefficient. As a consequence, current prototyping tools hamper iterative design practices. We present the design and implementation of RaPIDO a prototyping platform for this kind of interactive applications targeting interaction designers who have the ability to create interactive experiential prototypes but do not have electrical and software engineering expertise. A double case study is carried out to evaluate the platform in the context of a design process by following two designers developing outdoor games for children. We argue that RaPIDO allows a broader exploration of the design space for this kind of games and faster iterations. It allows designers to focus on the core of the game concepts they design rather than on complex and low-level engineering issues.

INTRODUCTION

In the last decade the field of mobile and pervasive gaming has evolved from creating early feasibility demonstrations to a vibrant sub-field of research and area of commercial product development. This interest is a natural consequence of the increased availability and miniaturization of mobile and pervasive technology. Here, we are particularly interested in technology-enhanced games where players are in the same location and where physical and embodied interaction is an integral part of the game play.

There are several types of games or more generally interactive play technologies that share these characteristics, e.g., Exertion games (Mueller et al., 2007), Open ended play (Bekker et al., 2010), Interactive playgrounds (Soler-Adillon and Parés, 2009) and Head Up Games. Exertion games are interactive games where players are induced to expend physical effort to play the game, with obvious health benefits for the player. They can even be played over a distance with remotely connected players. For example, the game *Breakout for Two* is a "cross between soccer, tennis and the computer game Breakout" (Mueller et al., 2003). Players are physically remote, and are connected via a life-sized video and audio connection; they can hear and see each other at all times. At each player's location a wall is used to project the other player on, enhanced with game elements (virtual blocks), generating a tennis-court like setup with the wall representing the boundary between the two players. The aim of the game is to make the virtual blocks disappear by hitting them with a (physical) ball. Though the players are physically separated, this setup encourages them to socially interact while engaging in a fun sport activity. Furthermore, Lindley et al. (2008) explore the impact of body movement in a computer game, by comparing the use of a standard videogame controller to an input device that affords natural movement while playing *Donkey Konga* and measuring the effect on the level of engagement and the degree of social interaction. Lindley et al. conclude that the amount of social interaction is higher for the input device that allows natural movements, and the findings suggest that in general the players become more expressive, which is interpreted as a sign of improved social interaction.

Supporting open-ended play with technology is a another approach that also aims to encourage for example physical or social activity, but is deliberately designed in an open-ended manner, without prescribing game goals or rules. Children use the interactive technology in a free and unstructured way and 'design' their own play experiences. For example, the *FlowSteps* (Valk et al., 2012) are a set of flexible, interactive mats, that children can engage with in an open-ended way: the children can position the mats to their liking and by stepping on them, the mats respond by lighting up in either red or blue. Children can explore the interactivity and develop their own games by giving meaning to the output modalities of the *FlowSteps*. Another example of interactive play objects designed to support open ended play are the *ColorFlares* (Bekker and Sturm, 2009): intelligent play objects that can detect movement (shaking, rolling) and respond by lighting up in one of six colors. Children are encouraged to explore the interactivity of the *ColorFlares* and attribute their own meaning to it to create playful experiences.

Research in interactive playgrounds addresses the lack of physical activity and social interaction that children these days experience by creating interactive installations on playgrounds that encourage full-body interaction. An example of such a setup is the *Interactive Slide* (Soler-Adillon and Parés, 2009). The slide consists of a large, inflatable, sliding surface that is augmented by projecting digital content on it. Several games have been implemented on it, and have been evaluated with users, showing its capability to engage children in full-body interaction. Less focused on physical activity, but more on the novelty of experience that an interactive playground can provide is the installation Breathless (Rennick Egglestone et al., 2011). Users are seated in an interactive swing, which they can control through their breathing.

Head Up Games are outdoor, interactive games that encourage physical activity and social interaction among collocated players, and are deliberately designed to give rise to play patterns traditionally characterizing children's outdoor games, like tag, and hide-and-seek. Typically, Head Up Games do not rely on display interaction, but instead on other types of interaction that do not interfere with the rich social interaction usually seen when children play outdoor games. *Save the Safe* (Soute et al., 2009) and *HeartBeat* (Magielse and Markopoulos, 2009) are examples of such Head Up Games.

Save the Safe | p. 36
HeartBeat | p. 37

The design of these games combines aspects of game design and interaction design, both of which typically benefit from fast iterative cycles, quick prototyping and play/user testing (Salen and Zimmerman, 2003). For example, the design of Head Up Games, e.g., *Camelot* (Verhaegh et al., 2006), *Save the Safe* and *HeartBeat* proceeded with traditional paper prototyping used to rapidly test the game rules. Next, low-tech prototyping of different aspects of interaction followed, and eventually a working prototype sufficient to support (parts of the) game play was developed. Especially the last step of moving from an early non-functional prototype to a working, interactive prototype is challenging, and requires, compared to the previous steps, a disproportional amount of time in the design process. As hardware configurations are purpose-built for the game, software typically needs to be written from scratch to deal with controlling anything from low-level hardware interrupts to game rules.

Camelot | p. 34

The situation outlined so far, naturally leads to the question whether or not recurring hardware and software design solutions can be re-used between games. In this chapter we describe a platform we have developed to facilitate prototyping the physical interactions needed to rapidly develop outdoor, interactive games. We focus on both the hardware, as well as putting effort in building a software library appropriate for designers not specifically trained in computer science, with the aim to lower the threshold for prototyping interactive play experiences.

We start by describing the challenges that designers face when creating interactive prototypes for games. Then we review existing platforms and toolkits that can support designers in the design process. Next we describe examples of characteristics and challenges a designer faces when designing Head Up Games. Based on these we present requirements for a platform. RaPIDO, the platform we designed, is presented and we discuss design considerations from a hardware, software and interaction design point of view. We then report two interpretive case studies in which we explored the value of RaPIDO as a prototyping tool in game design projects. We conclude by reflecting on the process of creating RaPIDO, both the platform itself and its evaluation and we discuss future work.

BACKGROUND

A generally accepted methodology for designing interactive products is the iterative design process. This is especially relevant in game design; Salen and Zimmerman (2003) emphasize the importance of playtesting games often and early in the design process. Costikyan (2002) supports this point of view; he argues that in the act of game design a designer aims to create a certain experience. However, the designer can manipulate merely the rules, goals and interactions of the game; the emerging game experience is ultimately created by the players themselves, which makes it virtually impossible to predict the player experience without actually playing the game. Slight changes to the game rules or the interaction can influence the experience and fun factor of a game in unforeseen ways. Ideally designers should design the games in short cycles, alternating game design and playtests rapidly to test effects of game interactions on the overall game experience.

To achieve this, especially in this new class of interactive games, designers need to make prototypes of their intended product. Well-known techniques to build prototypes in very early stages of the design process are paper-based prototypes and Wizard-of-Oz prototyping (Dahlbäck et al., 1993). Though these low-fi techniques can be useful in the early stages of Head Up Game design, as development progresses soon the need arises for more advanced, high-fidelity, prototypes, because the low-fi prototypes cannot provide the right timing of interactions to keep up with the high pace of the outdoor games. As a consequence the game under development cannot be properly evaluated.

Another motivation for building functioning prototypes early, stems from our own experience in designing outdoor games for children: we have observed that designers struggle making the step from a paper-based design to an interactive design because interactivity opens up a much broader design space in interactive games than what can be addressed by the paper-based prototypes. As a consequence, it appears that designing with low-fi prototyping media that cannot implement game rules, will steer designers to solutions that may be coherent, playable and fun, but for which the technology becomes a post-hoc addition that is not well integrated into the game play.

The mapping of design concepts to technological prototypes represents a non negligible challenge for designers. Especially when there is time pressure, compromises may need to be made that are driven by what is feasible for the design team to implement. It pays off to resolve some recurring challenges for designers, and to expose them early in the design process to the technology so they can assimilate possibilities and limitations, and even have more time to work with and work around technical challenges.

It is important here to make clear who the 'designers' are that we refer to. While the first examples of pervasive games might have been the product of technologically sophisticated teams of computer scientists, the evolution of the field means that increasingly interaction designers, versed in designing tangible and embodied experiences need to take on these technological challenges on their own. Currently the de facto standard for the broader area of tangible and embodied interaction, including wearables and information appliances, appears to be the Arduino platform* and related components (Hodges et al., 2012). Arduinos allow interaction designers with minimal electrical engineering skills to tinker together simple, interactive artifacts. It is this type of designer rather than the skilled software engineer that we address with this work. We assume that the primary interest of designers is the physical and embodied interaction and the game design rather than enabling embedded systems, where they currently often have to address recurring issues as power management, distributed control, parallelism, etc.

* http://www.arduino.cc/

In conclusion, it is preferable that the prototyping medium makes it easy to create experiential prototypes featuring the key elements of interactivity and game play. However, building a working, interactive prototype remains time consuming and expensive; requiring dedicated embedded hardware and software skills - skills a typical interaction designer might not possess to a sufficient extent. And though many rapid prototyping tools are available for the desktop environment, tools for rapid prototyping of physical interaction, especially mobile interaction, are less common. So, to better support interaction designers, specifically outdoor game designers, we argue that a tool is needed that offers a means to rapidly prototype outdoor gaming interactions (see also Figure 4.1). This will allow designers to design and test complex interactions earlier in the design process, thereby improving the overall interaction design and the emerging game experience. Furthermore, the platform can also be used as a tool for exploration of the design space; the field of outdoor interactive game design is relatively new, and we argue that for designers to tap the full potential of interactivity in outdoor games, they should have a quick way to play and engage with technology to get a clear design view how interactivity can be sensibly incorporated in games.

RELATED WORK

Though many dedicated hardware and/or software implementations exist for a range of outdoor, interactive games (see Magerkurth et al. (2005) for an overview), few attempts have been made to create generic and reusable hardware and software to support designers to work rapidly in an iterative design process. Here we review a set of relevant

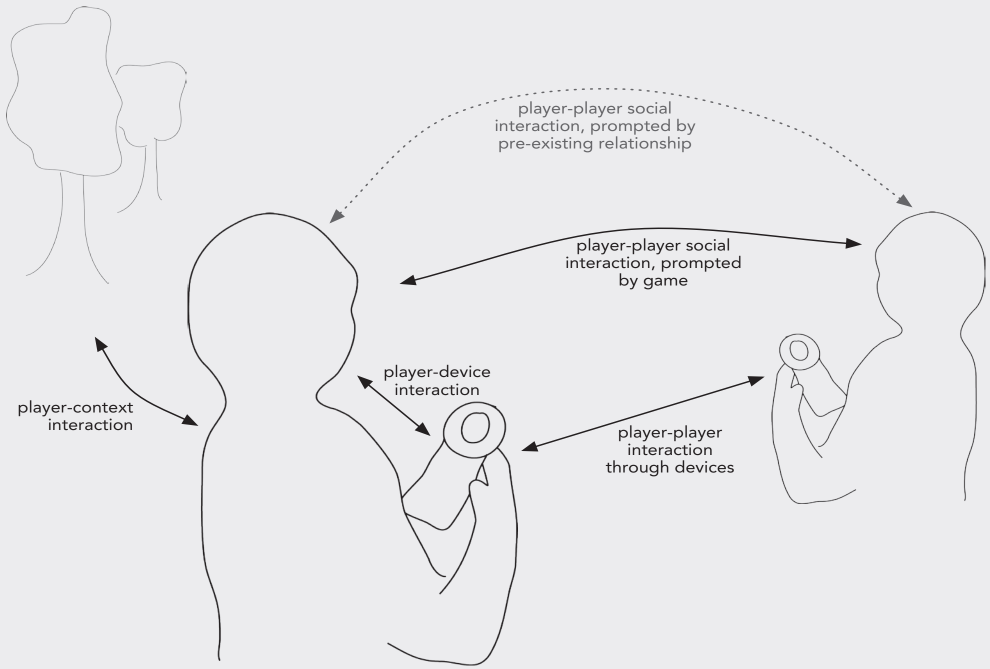


Figure 4.1 Types of interaction in that can be expected in Head Up Games. The evaluation of player-device interaction could be achieved in isolation with existing prototyping tools for multimodal interaction. However, in the case of Head Up Games many more interactions influence the overall game experience. Therefore it is important to evaluate the game designs in context and that the game devices support evaluation in such contexts.

tools and platforms for prototyping physical interaction. Roughly, these tools can be split in two categories: tools that support "sketching" (Buxton, 2007) interactive behavior and tools that support constructing interactive prototypes.

SKETCHING INTERACTIVITY

At the start of a design process designers commonly start sketching the intended product. However, though pen and paper are very useful to quickly sketch the form of a product, they are less able to convey interactivity. Several tools have been developed to combine the benefits of paper-based sketching with the merits of computer-based interactivity.

One of the first tools to merge sketching and interactivity was SILK (Landay and Myers, 2001) which can be used for quickly generating graphical user interfaces (GUIs) and designers can add behavior through storyboarding.

Similar, though not specifically targeted at creating GUIs, is idAnimate (Quevedo-Fernández and Martens, 2012): a multi-touch application for sketching animations that runs on for example an iPad. It allows designers to use gestures to create animations in a minimal amount of time.

Furthermore, Topiary (Li et al., 2004) is a desktop tool for prototyping location-enhanced applications. It allows designers to model locations, independent of sensing technology. Next, designers can sketch interface mock-ups, which can be "run" on a mobile device, in a Wizard of Oz-type study setup.

Finally, in contrast to the tools described above, Sketchify (Obrenovic and Martens, 2011) allows designers to integrate physical I/O in their sketches. Sketchify is a desktop program that supports I/O services for a wide range of input and output devices (e.g., Arduino, Phidgets) and also links to external development environments (Flash, Max/MSP).

TOOLKITS FOR BUILDING PHYSICAL INTERACTION

Several toolkits and platforms have been created to facilitate constructing prototypes that support physical interaction, with the aim to lower the barrier to prototype new concepts (Hodges et al., 2012).

Switcharoo (Avrahami and Hudson, 2002) is a prototyping tool that lets designers rapidly explore both form and interactivity of new products. The same makers present the Calder toolkit (Lee et al., 2004): a set of reusable input and output components that can be linked both wired as well as wireless to create an interactive physical interface. Both Switcharoo and the Calder system need to be tethered to a computer to control the interactivity.

Phidgets (Greenberg and Fitchett, 2001) are commercially available "building blocks" for easy composition of sensing and controlling technologies. A variety of components is available and also the Phidgets can be programmed using a wide selection of software languages and, like Switcheroo and the Calder system, Phidgets need a direct connection to a computer.

Similar to Phidgets are Arduinos: "Arduino is an open-source electronics prototyping platform based on flexible, easy-to-use hardware and software. It's intended for artists, designers, hobbyists, and anyone interested in creating interactive objects or environments." Arduinos have a microprocessor on board and offer multiple input and output channels to attach any type of sensor or actuator.

Specifically targeted at gaming is the VoodooIO Gaming Kit (Villar et al., 2007). VoodooIO offers gamers a way to appropriate their game controller towards a particular computer game. Attention has been paid to both the flexibility of form as well as the adaptability of the configuration of the game controller.

Switcharoo, the Calder toolkit, Phidgets and the VoodooIO Gaming kit have in common that they all require tethering to a computer for control of the interactivity. This makes them unsuitable for use in our context, namely in an outdoor environment. Arduinos can be used without a connection to a computer, as they have their own microprocessor to execute code. Designers, with limited electrical engineering and/or software skills, can relatively quickly put together a hardware and software combination to prototype simple, interactive behavior.

In the field of pervasive and ubiquitous computing several systems have been designed to develop applications: Magerkurth et al. present the Pegasus system architecture (Magerkurth et al., 2006), a component-based software architecture that enables designers to develop pervasive games while abstracting the actual interactions. This way, physical interactions and virtual interactions become easily interchangeable, allowing the designer to experiment with different configurations.

Furthermore, iStuff (Ballagas et al., 2003) offers lightweight wireless input and output devices to prototype user interfaces in ubiquitous computing environments. A graphical tool is provided so a designer can easily create event-mappings, i.e. linking input components (e.g., a light-switch) to output components (e.g., a light-bulb). The iStuff components require a connection to a server to connect to the smart environment.

In contrast to the former tools (Switcheroo, Phidgets, etc.) that focus on support for easy configuration of physical interactive prototypes, the latter tools (Pegasus and iStuff) focus primarily on addressing software challenges in ambient environments, such as: event-handling, management and coordination of multiple devices, autonomous detection of requirements and automatic service discovery and provisioning (Raychoudhury et al., 2012). Consequently, there is much emphasis on the (software) system architecture and the supporting infrastructure.

Considering the field we are addressing, outdoor, interactive games, we need a tool that is robust enough to take outdoors and lasts several iterations of game development. Furthermore, to truly evaluate the game experience, we need fully functioning prototypes; Wizard-of-Oz setups are generally too slow to keep up with the fast pace of the games we are creating. Finally, to support evaluations with interactive prototypes at the very beginning of the design process, we need a tool that has a very low threshold for engaging with it.

Though the sketching tools are designed to be used at the early stages of the design process, they can only be used to prototype the user-device interaction (see Figure 4.1). They are not suitable to be taken outdoors to evaluate the game experience in an outdoor context. Despite the fact that the "barrier to entry" has been lowered, the toolkits that support construction of interactive, physical devices still require some skills in electrical and software engineering. Moreover, if complex interaction mechanisms are required and several technologies have to be combined (e.g., radio, audio and RFID), the designer will face similar challenges to what engineers of embedded systems typically face, e.g., electronically integrating and interfacing the hardware components. Furthermore, it takes extra effort with to create robust prototypes that can be taken outdoors; typically, these toolkits rely on a physical link to a computer or external power source and can be fragile constructions as components tend to be loosely wired together.

We argue that the platform, which we present in a later section, is positioned between the sketching tools and the construction tools in terms of in ease of use and offered functionality; similar to the sketching tools it is usable very early in the design process, as it does not require time and effort to physically build the prototype. Similar to the creation tools, it offers an integrated set of working, interaction technologies. In contrast to the creation toolkits, our platform is specifically targeted at outdoor use, with an integrated power source and attention has been paid to creating a robust solution.

Before presenting the platform, in the next section we will first elaborate on scenarios in which such a platform could be used.

EXAMPLES OF USE

In this section we will revisit three Head Up Games to illustrate the typical design space and context we are addressing and the challenges we are facing. The examples used are games we have designed in previous projects, though we argue that issues we have encountered are generalizable for other outdoor games as well. We shortly mention the specific technology and challenges for each game here, for more elaborate information, please refer to the earlier descriptions of the games.

CAMELOT

Camelot (Verhaegh et al., 2006) was designed by a group of three designers during three months.

Camelot | p. 34

Technology *Camelot*'s technology consisted of LEDS wired to a PIC processor. Communication between devices was implemented using infrared technology. The software was written in JAL (a basic, textual programming language).

Challenges Initially we had trouble deciding which technology to use to identify a zone as our budget was limited. Also, due to time constraints not all elements of the game could be implemented in technology; for example, though the pick up of the resources at each zone was implemented, the drop off at the castle was not and had to be prototyped using a Wizard of Oz setup.

STOP THE BOMB AND SAVE THE SAFE

Stop the Bomb (Hendrix et al., 2008) was designed by four design students during three months; subsequently *Stop the Bomb* was adapted by one single designer to create *Save the Safe* (Soute et al., 2009).

Save the Safe | p. 36

Technology For *Stop the Bomb* and *Save the Safe* we used Crossbow motes*, which were commercially available at that time. Crossbow motes automatically form a wireless sensor network, mostly used for monitoring purposes (e.g., to monitor the ambient temperature in a warehouse). For our games we attached several LEDs and a vibration motor to the motes. The motes were programmed using C.

* http://en.wikipedia.org/wiki/Crossbow_Technology

Challenges In the given time we were unable to write in software a robust algorithm for randomly assigning players to a team. In the end, we settled for a setup where a crossbow mote that was not involved in the game, was connected to a laptop. Through this connection team assignments could be transmitted from the laptop to each player's crossbow mote, simulating the random team assignment.

Another challenge was the communication protocol for sending messages. Though the motes automatically form a sensor network, they did not provide a protocol for asserting that messages arrived at the receiving mote. A crude 'handshaking' protocol was devised that entailed rapidly sending back and forth a few messages to be reasonably sure that the message had arrived.

HEARTBEAT

HeartBeat | p. 37

HeartBeat (Magielse and Markopoulos, 2009) was designed in a project by an industrial design student over the course of six months.

Technology PIC processors were used for *HeartBeat*, combined with an XBee module for wireless transmission of messages between devices. Furthermore, an off the shelf heartbeat monitor was used to measure the player's heartbeat.

Challenges Similar as in the *Save the Safe* game, in this game the random assignment of players to a team posed problems. Also integrating the heart rate monitor with the rest of the hardware was not straightforward.

COMMON CHARACTERISTICS AND CHALLENGES

Each game employs several interactive game devices, mostly one for each player, or alternatively shared by a team. These devices were small and portable to be easily carried by players during game play.

On an abstract level common functionality can be distinguished, for example the need to identify other game objects, or the need to send message to other game objects. However, what the examples above illustrate is that very different approaches for implementation in hardware and software were chosen, each time requiring to go through the process of selecting and appropriating the specific hardware and software.

All examples have in common that during implementation we experienced time pressure; there was always too little time for implementation. Though the initial design phases of these projects in general went smoothly, once implementation started progress would slow down considerably, as the designers in these projects were interaction designers, and not trained in the field of embedded systems engineering. This resulted in completing only one design cycle with (partly) functioning technology. Ideally we would have wanted to complete more cycles for iteratively improving the game design.

Finally, it is well known in the field of human-computer interaction that when major effort is invested in engineering a solution there is likely to be a reluctance on the part of its creator to take on board user feedback during evaluations and make necessary changes. It becomes essential then to lower the cost of prototyping (Virzi, 1989) .

PLATFORM REQUIREMENTS

In the previous section we have listed common challenges and characteristics of three Head Up Games. Based on this list we now propose the requirements for a platform that must support rapid development of outdoor, interactive games.

Connectivity

Connectivity between game devices is essential for the game play. The ability to relay game events between devices provides a designer with a considerable amount of flexibility in the game design in contrast to stand- alone game devices. Though we argue that connectivity is essential, we also argue that this could add to the complexity of the total system, e.g., if adherence to standard network protocols are necessary. We argue that for the type of games we are creating, peer-to-peer connections are sufficient, as the examples in the previous section illustrate. Pervasive applications requiring the connection to a network, are outside the scope of the present platform, and are supported by prototyping platforms such as (Ballagas et al., 2003; Magerkurth et al., 2006) described above.

Distributed

In our vision of Head Up Games (see Chapter 2) we stated that game play should not be dependent on the existence of infrastructure (like GPS coverage or availability of a Wi- Fi network), similar to traditional toys as a ball or hoop that children can pick up and play with anywhere. The games illustrated in the previous section show that GPS and/or WiFi are not a necessity for outdoor, interactive games. So, the platform needs to be self-contained, both with respect to powering it as well as controlling it.

Robustness and portability

As the platform is aimed for use in early stages of the design process, the prototypes need not be as robust as a final product would need to be, but still the prototypes will need to endure a 'harsh' environment. In contrast to prototypes that are tested in a 'safe' environment as an indoor lab setting, these prototypes are taken outdoors and will be quite roughly handled as players are running about. Also, they need to be lightweight and small enough to run around with, without interfering with the physical activity that is common in outdoor games.

Interactions and extensibility

Interactions that are appropriate in outdoor games include audio and visual feedback to signal game events. Children often perceived haptic feedback as fun, as was shown in *Save the Safe*. Not all types of interactions can be foreseen beforehand and included in the design of the platform. Therefore it is necessary that the platform is extensible so that additional sensors and/or actuators can be connected. Further, consistent with the type of interactions (embodied interaction) we wish to support, we do not aim to support text input or graphical screen-based interaction.

Easy to program

Interaction designers are generally not trained as computer scientists; they are not formally educated in designing and implementing software. However, most designers are technically inclined and have some experience in programming for example Flash or JAVA. Thus, the software of the platform should support the designers' skill level.

Transparency of hardware control

Low-level coding of the hardware components can take up a significant amount of time, so, the platform should offer a higher-level API for the designers to easily control the hardware.

Reusable interaction styles across games

Though the actual game play of different games can vary greatly, still common features can be identified that return in most games, for example assigning players to a team, or detecting nearby players. So far, we have implemented this type of functionality each and every time - though designing these kinds of algorithms is not so trivial for nonprofessional programmers. The platform should therefore not only offer an API to address the actual hardware, but also offer a set of functions that abstract common game features.

PLATFORM ARCHITECTURE

In this section we describe the platform RaPIDO (Rapid prototyping of Physical Interaction Design for Outdoor games) we have created that meets the requirements as described in the previous section. Throughout the text we refer to 'the platform' (or RaPIDO) and 'devices' or 'prototypes'. To clarify: with 'the platform' we mean the general architecture of hardware implementation and software libraries to control the hardware; one instance of the platform is called a device or prototype. To prototype a game, designers can use one or multiple of these identical devices to build for example player's devices and/or other interactive game objects.

INTERACTION ARCHITECTURE

Based on our experiences with previous Head Up Games we defined a set of interaction styles that we found necessary for Head Up Games. Common interaction styles in Head Up Games are: detecting objects, measuring the distance to other players, receiving direct input (e.g., using a knob or detecting a shake), communication with other devices, giving auditory, visual and tactile feedback. Table 4.1 lists the main components that support these interactions.

Leaving out a display was also a deliberate choice. Obviously, in Head Up Gaming we want to promote rich social interaction and demote the continuous use of a display as this interferes with Head Up play. To make a clear design statement, we have chosen to leave a display out of the design. This basically comes down to a matter of affordances; we expected that the mere availability of a display would prompt designers towards traditional screen based graphical interaction, which is outside of our intended

Table 4.1 Main components of RaPIDO

Technology	*Interaction style*
4 RGB LEDs	provide visual cues, e.g., by blinking, or changing color
Sound chip + speaker (SD Card)	provide auditory cues, can read and playback .wav files from SD card
RFID module	detect objects tagged with RFID-tags
XBEE module	provides: (1) inter-device communication (2) distance measurement between devices
Vibration motor	provides tactile feedback
Rotation encoder	measures degree of rotation of wheel
Accelerometer	measures movements

scope. Though we did not explicitly test this hypothesis, it can be implicitly deduced from the types of games that have been presented so far in related work, see for example Benford et al. (2005). These games were all developed on PDAs and smart phones and although other technologies were available for interaction, still the designers extensively used the screen to interface with the players, inevitably leaving less room for rich interaction similar to traditional outdoor games. This convinced us to purposefully exclude a display from our design.

HARDWARE ARCHITECTURE

Not directly linked to supporting interaction styles, though equally important are the general hardware choices for, for example, battery and processor, as they affect the mobility and portability of the devices. Each device must be able to function stand-alone, so it needs its own power supply and its own controlling mechanism, without needing an extra computer that controls all devices remotely. The battery has been chosen by balancing the trade-off between dimension and power. Furthermore, we selected the Arduino Mega* microprocessor to achieve stand-alone operation: designers can write code on a computer first and upload this to the device after which the device can execute the code independently. The reason for selecting the Arduino board is that it is a well-known platform for supporting the creation of experiential prototypes for tangible and embodied interaction and on the Internet there is an active community that provides support.

* http://www.arduino.cc/

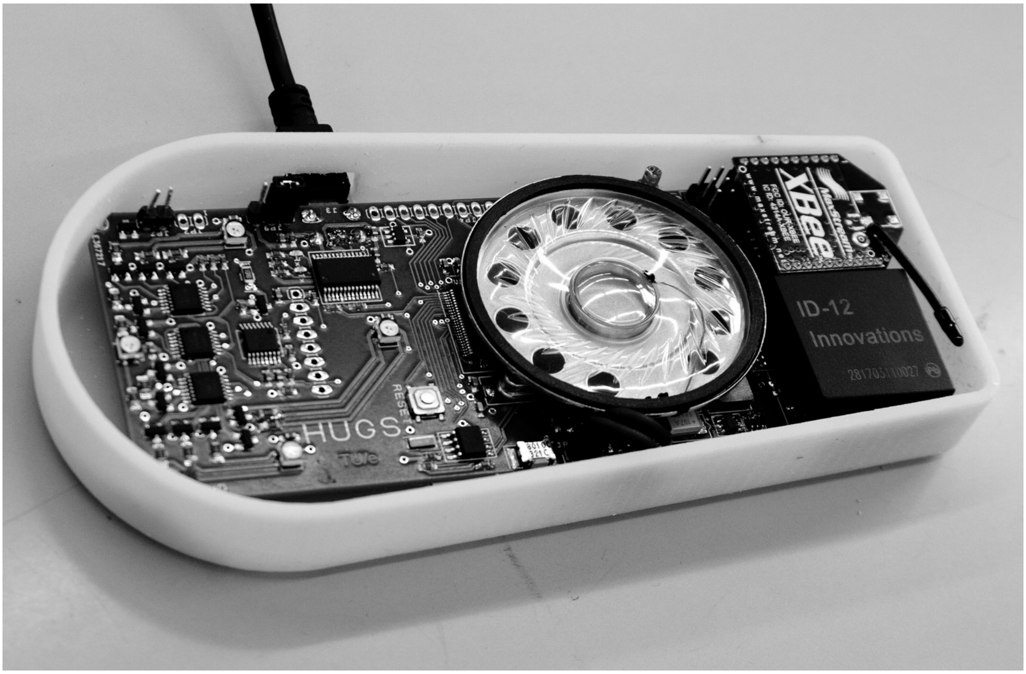


Figure 4.2 Impression of the hardware

As the standard Arduino Mega board was too big for our needs, we designed a dedicated printed circuit board (PCB) for our prototypes (see Figure 4.2). Not only did this result in a smaller design, it also improved the robustness of our devices: normally, when building applications on the Arduino boards all components are separately wired to the board. Because we needed to integrate many different components, this would have led to a complex manufacturing process and would have resulted in a less robust prototype because wires can come loose.

Some design choices address key requirements for the platform. For example, a GPS chip was not integrated; we found that in the games we have designed so far, there was never a need to get an absolute measure of the location of a device. We did find it useful to have a measure of the relative distance between devices, but for that a GPS is not necessary; we simply use the signal strength of the radio signal to calculate relative distances (as was the case with *Save the Safe*, and *Stop the Bomb*). Though less accurate than GPS it proved to be reliable and fast enough for Head Up Games.

Implementation trade offs

In the process of implementing the devices inevitably tradeoffs needed to be made to address pragmatic constraints. The main factors enforcing trade offs are the size and feature set of the device and the total costs. These trade offs are shortly discussed here.

We have put much effort into selecting an appropriate battery. As the device needed to be portable for children, its size and weight mattered. On the other hand though, we needed a considerable amount of energy to power all the components on board, ideally requiring a large battery. We settled for a LiPo battery (similar to a mobile phone battery). Furthermore, all components are carefully selected considering costs, functionality, and battery consumption.

The platform is based on an Arduino Mega processor and we designed our own dedicated PCB. Ideally, we would have designed a very small PCB, which is technically feasible. However, the costs of manufacturing a PCB are directly related to the size of the PCB and the size of its components. Also, another important consideration with respect to size is the flexibility for making changes post-production. We assumed that the first series of PCBs would inevitably contain flaws, and it is much easier and cheaper to use larger components that can be replaced with standard equipment rather than very small components that can only be replaced using dedicated, costly, equipment. So, we balanced size, costs and flexibility, resulting in the size of the devices now being slightly larger than we intended it to be initially, but in return affordable cost-wise and manageable from a technical point of view.

SOFTWARE ARCHITECTURE

As stated in the Requirements Section the software of the platform needs to match the general level of skills of interaction designers. We decided to adopt one of the most widespread and supported programming languages available: C.

By selecting C we argue that we are able to serve our target user group best, because it is possible to adapt the complexity of the software depending on the skill level of the designer. First, a designer is not forced to use a dedicated programming environment: for C many programming environments are available. Thus, one can pick the programming environment that most suits ones needs and capabilities. Novice

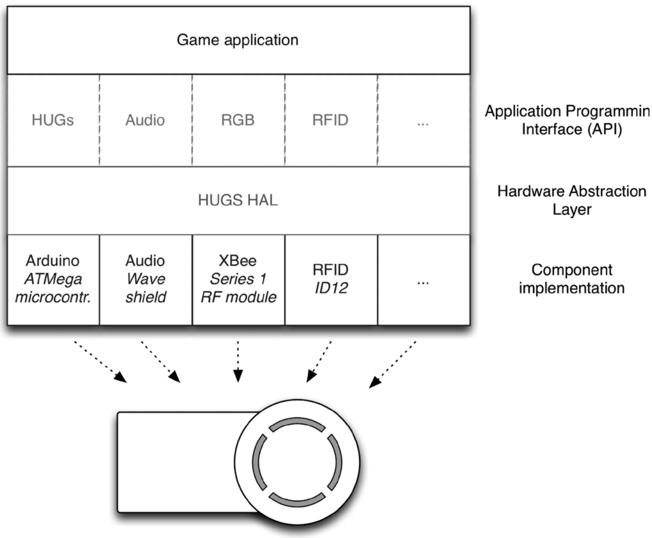


Figure 4.3 Overview of the general software architecture

programmers might choose the native Arduino environment; the API libraries of the platform can simply be imported and used with the Arduino language. The Arduino environment is easy to set up and run, but limited in functionality. Alternatively, more skilled programmers might choose an advanced environment like Eclipse*.

* http://www.eclipse.org/

Second, the software for the platform is designed in layers, in such a way that novice programmers can use the most simple interface, and advanced users, who might need more control over the platform, can dig deeper in the software to suit their needs.

Finally, by providing designers with a software library of common functionality development time is shortened and reuse is encouraged.

Figure 4.3 shows the general architecture of the software. First, directly linked to the hardware, a layer of dedicated software libraries was built in C++, using an object-oriented approach. Each library addresses a particular component of the hardware and contains low-level details of the particular hardware used, e.g., to what digital or analog ports the hardware is physically wired to the microprocessor. These libraries are strictly separated and are not dependent on each other. This way, if ever one of the hardware components needs to be replaced by another brand or make, only the corresponding software library needs to be adapted, without affecting the rest of the software.

The component libraries are used by the hardware abstraction layer (HAL). The HAL manages the creation of all objects in the software, ensuring that the number of objects instantiated in software directly matches the components in the actual hardware. Furthermore, the HAL offers general functionality for example a scheduler, and battery management functionality. Like the component libraries, the HAL is not accessed directly by the end-users.

Finally, the application-programming interface (API) is built on top of the HAL. These are the libraries that are typically used by designers building software for their games. In contrast to the component libraries, these libraries offer task-specific functionality, i.e. game related instead of technology centered. Engebretson and Wiedenbeck (2002) argue that this way it is easier for novice programmers to code their software. For example, in previous designs of Head Up Games we saw that sending a message to a nearby device is often what designers want to achieve. So the API offers a function `sendMessageToClosestDevice()`, which linguistically is a function name close to what a designer most probably wants to achieve when programming his game. The API handles the detection of the closest device and subsequently will send the message to that device.

Furthermore, these API libraries sometimes do have interdependencies, for ease of use for the designers. For example, one of the functions offered by the Head Up Games library is `assignTeam()` to dynamically, at run time, assign devices to a team using RFID tags. In his code the designer simply specifies which RFID tag and color are related to what team, and then the rest is handled by the Head Up Games module. Of course, designers in our target user group could build this functionality themselves by calling separate functions in the RFID library and the RGB library, but since it is a recurring function in many games it pays off to have a readily available and robust implementation.

DESIGN OF THE PHYSICAL SHAPE AND FORM

In related literature it is often mentioned that there exists a gap between designing the interaction and designing the form of a product (Avrahami and Hudson, 2002). More often than not, the form of a product is independently tested from the interaction of a product, though the test results would benefit much more from testing both simultaneously. Several of the toolkits presented in the Related Work section try to address this issue by offering flexible materials that designers can shape to their liking.

For our platform we decided not to offer a dedicated material to shape the casing of the prototypes. The functioning of the prototypes does not rely on the material of the case either, so instead designers can choose to apply any material they prefer to design their own case. Nowadays, 3D printing has become more accessible to designers (e.g., Shapeways*), so it is entirely possible for designers to create their own designs around the hardware relatively fast.

* http://www.shapeways.com/

Nevertheless, if designers do not (yet) want to explore the form of their design, but instead quickly want to explore interactions, we have created a dedicated case for the prototypes (see Figure 4.4). This case was designed to perfectly fit the hardware and extra attention has been paid to durability and robustness.

EVALUATION OF THE PLATFORM

In the previous sections we have described the creation of the platform and the design decisions we took along the way. In this section we will describe the study we have run to verify whether the platform was suitable for use by designers when prototyping games.

GOAL

The goal of the study is to answer the following set of questions:

Concerning the design process

- Does the platform support designers to rapidly prototype their concepts?
- Does it affect their design process, and more importantly, if so, how?

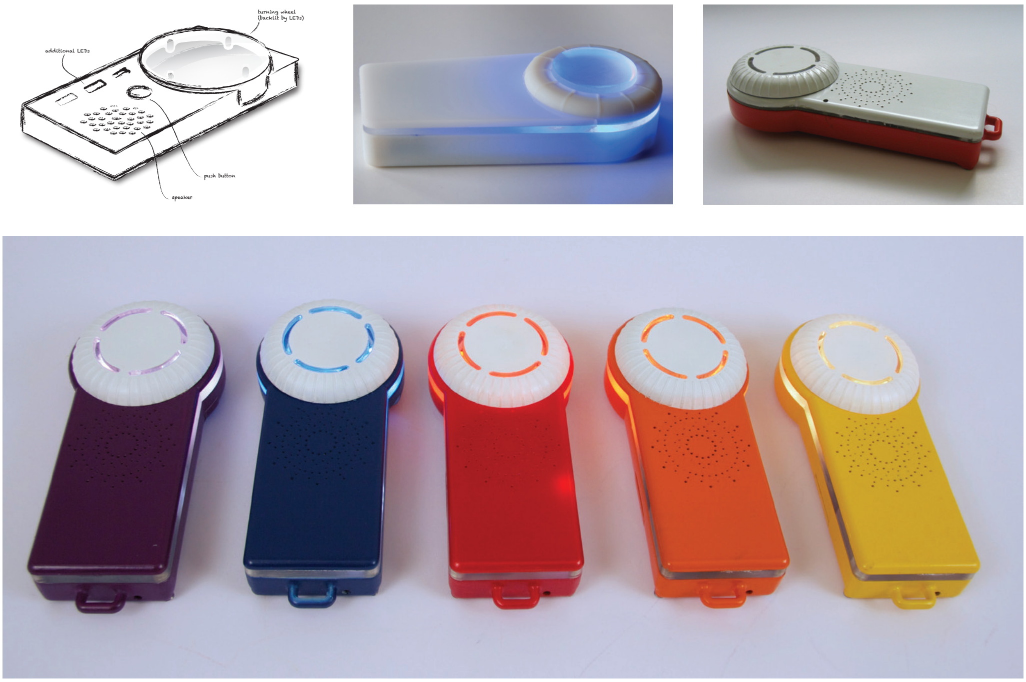
Concerning the functionality of the platform

- Does the platform offer suitable hardware and software functionality to create mobile interactive games/ applications?
- Identify areas for improvement of the platform (e.g., usability issues)

To answer these questions we decided to conduct a longitudinal, informative case study (Klein and Myers, 1999), following two designers for several weeks. Especially answering the questions about the design process can only be properly answered by following the process over a longer period of time rather than observing the execution of preset programming tasks which is the approach often taken for the user evaluation of prototyping platforms or software APIs. We argue that as the design and the design process evolve over time, new insights with regard to the platform will influence the game design and new insights on the game design will have an effect on the way the platform is used. But also matters like usability issues are arguably better to identify in a longitudinal setup: as the platform is new for the participants, there will be a learning curve for the participants that cannot be accounted for in a short evaluation session.

Thus, we decided against given the participants pre-set tasks, but instead asked them to use the platform in the projects they were working on in a real context. There are a few reasons to do this: first, we argue that it is important to let the platform be used within a real context, to ensure ecological validity, especially with respect to the questions

Figure 4.4 From first sketch (top left), through two iterations of prototypes (top right), to final prototypes (bottom)



relating to the design process. Second, in this study we are asking a considerable amount of input and effort from the participants; by letting them use the platform within their own design projects, participation in the study creates value for them as well.

SETUP OF STUDY

Our research question is twofold: we want to track the design process of the participants and see whether the platform has an influence on the process. At the same time we want to identify issues and areas for improvement for the platform.

To gather insight in how the participants perceived the platform and how their perception changed over time, we chose to adopt the concept maps method as proposed by Gerken et al. (2011). The method is developed to "elicit the mental model of a programmer when using an API and thereby identify usability issues and learning barriers and their development over time". The main idea is to let participants construct a concept map that reflects their understanding of how the API works in relation to the participants' own software. Furthermore, participants are asked to attach a set of adjectives to concepts on the map - the adjectives, but even more the changes over time in the placement of the adjectives, can indicate usability issues with the API. Each session, the participants are asked to modify and extend the concept map (including adjectives); allowing researchers to study the changes in the concept map and thereby identify issues with the API. In contrast to Gerken's study we are dealing with non-professional programmers, who are not only using an API, but are also working with the corresponding embedded hardware. That aside, we expect that using the concept maps method is an appropriate way to identify usability and learning issues and track changes in the design process over time.

To be able to track the evolution of the designs, we asked the participants to create a pitch of their intended design; a short written description of the game. Similar to the creation of the concept map, we planned to ask the participants to revisit the pitch every session, updating areas that were (re-) designed and/or adding more details as these became apparent.

In the next section the participants are described in depth, as their background and skills are relevant in the context of using the platform. Next, we describe the study protocol that we followed.

PARTICIPANTS

For this study we followed two participants. They were selected as their background and skills represent the target user group of our platform: interaction designers who want to design and rapidly create games for their studies. They are educated in industrial and/or interaction design, did not receive formal training in software or embedded engineering, but are able of writing simple code. The two participants we selected are described below.

Sofia. Sofia is a master student working on her final project to obtain her Master's degree at the Industrial Design department of our university. In a previous project, Sofia has worked to create games for children, though in that project she did not progress past the conceptual phase. Sofia has limited software skills; in other projects she has used JAVA, mainly to create mock-ups of web-based applications. Typically, these projects lasted one to two weeks, with various degrees of user involvement. She has no experience with embedded technology.

Sofia plans to use the platform to rapidly create several outdoor games for children. For the purpose of this study, we will only be tracking the design process of one of these games, called *Follow the Light*. She plans to develop and test the games in an iterative design process, involving children as soon as possible.

Follow the Light | p. 68

Ron. Ron is a master student at the Industrial Design department of our university. He is currently working on his final project to obtain his Master's degree. Though he has followed no formal software education, he has experience in writing software for Arduino and Processing.

Ron has done previous projects in which he designed interactive technology for users, and has created working prototypes. Emphasis has been on the design and the interaction on a conceptual level. However, he has conducted limited evaluations with end users during the process of designing, be it as informants before designing a concept, or as evaluators once the concept has been created.

Mostly, because the implementation of the technology took a considerable amount of time, there would be no time left to evaluate the product.

Ron wants to use the platform to implement an interactive application for children, called Manitou. In Manitou each child is equipped with an amulet that represents an animal. Each animal has certain characteristics and personality traits, and when the amulets come within each other's reach, situations develop in which the animals' moods are influenced by each other's presence.

STUDY PROTOCOL

Note that all meetings with the participants were held individually.

Initial meeting

At the start of the study we interviewed the participants to gather insight in how they normally engage in a design process. We specifically asked to reflect on previous projects, and not on their ideal process, to be able to identify more realistically later how using the platform has affected their design process. Furthermore, we asked the participants to create a pitch of their design: a written statement of their concept.

Follow-up meetings

We met the participants every few days; the time between meetings was dependent on how much time they had spent with the platform. For example, if the participant had spent a week of full time work with the platform, we would meet twice that week. If (s)he had worked less, one meeting was sufficient.

During the meeting, which was audio recorded, we followed the same protocol:

1. Creation/adaptation of the concept map. We asked the participants to think aloud (Boren and Ramey, 2000) while doing this, though we did not follow a strict think aloud protocol; during the concept map making we asked for details or posed questions that arose.

    - At the first follow-up meeting we asked the participant to create a concept map of their current understanding of the platform by sticking post-its (the concepts) on an a sheet of paper and drawing relationships between their own concepts and the platform concepts. We provided the participants with platform concepts (in green post-its) and with concepts from their own pitches (in yellow post-its). Participants were allowed to add concepts as they saw fit. In contrast to Gerken et al. (2011) who let participants extend and modify the existing concept map, we asked participants in subsequent meetings to create a new concept map. For the participants to be able to reflect on changes since the last meeting, we showed them their previous concept map.

- We asked the participants to rate (some of) the concepts using a set of bipolar constructs (adjectives): good - bad, beautiful - ugly, difficult - easy, clumsy - elegant. A concept could be rated with only one construct at a time, and we did not enforce the use of all adjectives; we asked the participants to only apply the adjectives they thought relevant at that time.
- We asked the participant to identify problem areas by drawing, with a red pen, a line around that particular area of the concept map.

2. We asked the participant to read out loud their game pitch, reflecting on changes that had occurred since the last meeting.
3. We conducted a semi-structured interview, asking about their progress with their design. The participants could ask detailed questions about the platform.

In between contact

We instructed the participants that, if they ran into issues with the platform, they could contact us at any time in between meetings, by any means that they preferred (e-mail, phone, drop by in person). Though the participants could of course report issues at the regular meetings, we need to acknowledge here that both participants are working on their final projects, and time is an issue for them. Also, these issues could be resulting from the platform either not working or participants misinterpreting the use of the platform and therefore are important data for us to capture.

The role of the researcher

The platform is meant to rapidly develop mobile, interactive games. We ourselves are knowledgeable in the area of game design, more specifically with children. However, to not influence the process that we are studying here (i.e. the design process of the game designers and how the platform supports this), we decided that we would refrain from commenting or giving advice on the game designs. Thus, we restricted ourselves to answer any technical questions that would help the participants implement their game on the platform. This included explanations on how the hardware worked; explanations on software functions, referring them to functions they had not identified as useful; etc.

Furthermore, in our dealings with the participants we made a conscious effort to create an open atmosphere, in which the participants would feel confident to voice their negative findings of the platform too, instead of feeling obliged to give pleasing comments only.

Materials

Both participants received two devices to start developing their games with. In total we had ten devices available, which they could obtain on request after agreeing with each other, so as not to interfere with each other's development and evaluation plans.

Furthermore, they received a manual of the platform (see Appendix B), describing the hardware and software, including examples on basic use of all the components. The software furthermore contained elaborate comments in all header files, explaining each function and a Wiki style website was available with information on the software. We helped them setup their computers for developing the software in their preferred environment (either the Eclipse IDE or the Arduino IDE, on Windows or Mac).

EXPECTATIONS

We have several expectations regarding the adoption of the platform in the design process; we present them here to be able to reflect on them later:

- We expect that, during the process of adopting the RaPIDO platform, the participants will the need to adapt their initial concept. This can be for two reasons: 1) They find out that something (in hardware or software) is possible, which they did not anticipate, and decide to integrate that into their design. 2) They find out that something is *not* possible, and either remove that from their design, or

redesign that particular feature by switching to another hardware/software strategy. For example, participants might expect to use GPS to measure (absolute) distances between players. However, GPS is not incorporated in the platform. Therefor, they decide to switch to using the signal strenght measurement of the radio to measure relative distances.
- Industrial design students are not trained in computer science. Often they are quite capable of creating short pieces of code; however, when the complexity of the interaction is increased, we expect that they will have trouble managing the complexity of their code.
- We expect the students to be able to create a first, functional prototype that is ready for testing within two weeks (full time work), i.e. at least one modality (e.g., the sound system) is fully implemented in the game.

RESULTS

Sofia developed several Head Up Games in an iterative process; the results of that process have already been reported in Chapter 3. For the purpose of this study we followed the development of the game *Follow the Light*. It took Sofia 2½ weeks to go from the concept to working prototype; in total we held four meetings with her.

Follow the Light | p. 68

In total we held six meetings with Ron, over the course of two months. Within 3 weeks he had implemented working test-code for each of the components he was planning to use for his game. After that his use of the platform itself remained quite stable, his focus was totally on writing the game-logic.

During the evaluation period, we made some changes to the software of the platform. Most notable is the software for the radio component. While testing Ron uncovered an until that time undetected technical problem, which we needed to fix. Furthermore, both Ron and Sofia requested functionality that was initially not provided by the platform. If it was relatively easy for us to implement, we did so; otherwise, we helped them to adapt their software design as to achieve their desired end result. For example, the API offered functionality to send a message to the closest device. In effect, the API shielded the designer of first finding out which device is closest and subsequently sending a message to that device. However, for his game Ron wanted to get distance information between devices; information that was initially not passed on to the user of the API. It was very easy to make this information available through the API, which is what we did.

No changes were made to the hardware.

ANSWERING THE RESEARCH QUESTIONS

From each session we gathered data in the form of maps (with the adjectives and 'red lines'), pitches and interview transcriptions. Furthermore, we kept logs of any questions posed in between meetings.

We analyzed the data by looking in detail at each session, in particular at changes that were made with regard to the previous session. These changes could indicate many things: a change in the design process, a new step in the learning process, a changed perception on how the platform works, a usability issue. From the interview data we were mostly able to attribute changes in the maps to any of these reasons. Furthermore, we looked at the maps and pitches in isolation, to identify possible trends over time. Also, we regarded the design process of the participant as a whole and compared it to their previous design processes. Finally, we discussed all our observations with the participants at the end of the study.

With all this evidence gathered we attempted to identify clues that (possibly) answered our research questions, on which we elaborate below.

Does the platform support designers to rapidly prototype their concepts? Does using the platform affect the design process and, if so, how?

We argue we have found evidence that supports this: both Sofia and Ron were able to generate working, interactive prototypes, within two to three weeks, which is inline with our expectations. Sofia indicated that by herself she would never have succeeded creating the hardware; Ron said he would have been able to do so, but it would have cost him much more time to select the right hardware components and assemble them. With the platform they both seemed to be able to better focus on the actual design of their applications, and less on the details of the embedded hardware. Furthermore, we observed that the platform allowed the participants to freely explore the technology,

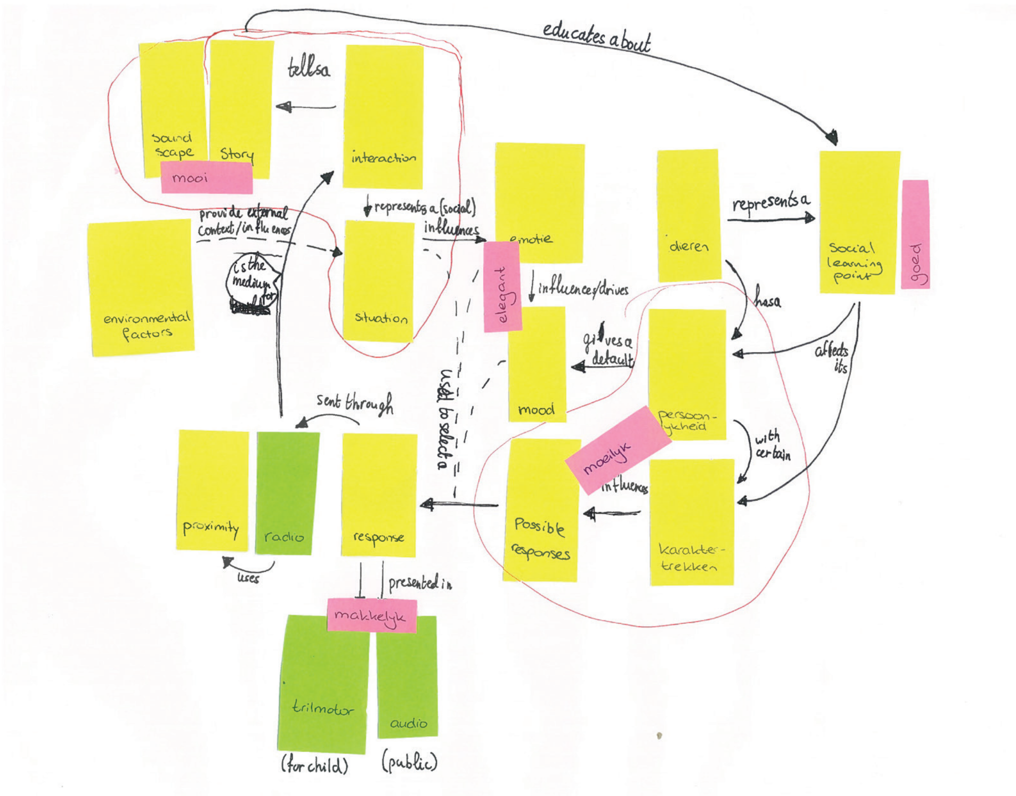


Figure 4.5 Final concept map Ron, showing that the focus of his work is on his game concepts (yellow post-its) and less on the platform concepts (green post-its).

which influenced the design process positively, allowing them to quickly and effectively incorporate technology affordances in the design concept.

In both Ron's and Sofia's case we observed that the readily available technology helped them focus on the game interactions. For example, Figure 4.5 shows Ron's final concept map: green post-its represent platform concepts, yellow post-its represent Ron's own game concepts; clearly, the latter are more present and have more complex connections with each other. This indicates a clear focus on his own game software. Furthermore, when asked if he thought whether or not the platform supported him, Ron answered: *"Yes, it definitely helped me. For the application that I am building [the platform] essentially provides the backbone, the part of the hardware I normally create first. That has been done for me now. The more complex implementation of my game, that is something that can be based on [the platform]. Now I have access to [the platform], [...] I can focus on my own design."*

In Sofia's case we noted that when she was making her first concept map, she pondered on the several concepts of the platform, rejecting the concepts "battery" and "arduino (processor)". She argued that she did not need them. Of course, the devices would never work without battery and processor, but the fact that she perceived these components as unnecessary, indicates that she has a focus on the game design and that the platform relieves the designer of dealing with lower level embedded components.

We observed that the platform allowed the participants to truly and freely explore the technology, i.e. try out several different interactions quickly and immediately see the effect on the game play. Also, by easily engaging with the technology they formed a better picture of how it worked and thus could better judge how to implement their concepts. Finally, having the technology available meant that the game design and the technology implementation were developed concurrently, each affecting the other. In contrast to previous designs (e.g., *Camelot* (Verhaegh et al., 2006)) where the game design would have been first more elaborately detailed on paper before proceeding to a technology implementation. For example, in her game Sofia needed the devices to express a particular value to the players. In her first concept she intended to do this by letting the LEDs flash a number of times, indicating the value. However, while she was trying things out, it occurred to her that this would not work in the real game; it was not certain that players would be looking at the devices as it was flashing, and after the flashing stopped, the "information" was gone. Subsequently she redesigned that particular part of the interaction to a more persistent way of showing the information. She reflected on this that it was easy to find out, as the platform allowed her to quickly try things; would it have cost a considerable amount of time, she would have probably not done this and would have found out during a user evaluation.

Camelot | p. 34

Does the platform offer suitable hardware and software functionality to create mobile interactive games (applications)? How can the platform be improved?

By having the participants use the platform in their design processes we are able to reflect on the suitability of the platform for mobile, interactive applications. Also, we are able to identify areas for improvement.

The hardware

Both Sofia and Ron indicated that the hardware was sufficient for their needs. The radio was very often used in the game designs, to transfer game-related messages between devices and also to measure distances between players. Furthermore, both audio and the RGB LEDs were used in most games to give the players feedback. Finaly, Sofia made heavy use of the RFID reader; sometimes for game-related interactions, but often to set up a game. For example, she would load three games simultaniously on the devices and program three RFID tags, one for each game, to be able to select one of the games.

At one point, Ron indicated that he would like to have some way of detecting the location of the devices. Because a GPS chip was not integrated on the RaPIDO platform he initially had no idea how to deal with this and requested our help. In discussion with him we came to the conclusion that in fact a distance measurement would be sufficient for his game, which is information that is possible to retrieve from the hardware.

In his map Ron indicated that he thought the form of the devices was "ugly". He had already done a few form studies of the look and feel of the end product, which did not resemble the current form of the RaPIDO devices. Eventually, he would want to discard the current casing of the devices and integrate the platform's printed circuit board in his own design; however, he remarked that the size of the PCB was still a bit too large for that (note that handcrafting electronics using the same components, would probably result in even larger-sized hardware).

The software

We observed that both participants had very little trouble implementing the more direct interactions in their design, for example playing an audio file, reading a value from the rotation encoder, or combining these interactions. They had no trouble identifying the correct functions to enable these interactions.

In the beginning, Sofia experienced problems with the C syntax, e.g., creating a for-loop, but these problems declined after about a week. Ron had previous experience with C and seemed to have no problems at all. However, we found out that both participants struggled with more complex syntax in C. For example, the library contained a scheduler that could schedule the execution of a function at a later time. Though that was found useful, at first Sofia did not understand how to use it; the function for setting up a task in the scheduler accepted a function pointer as an argument, and neither participant had previous knowledge of this concept. Once we explained the syntax of the function pointer, Sofia started to use the scheduler in her designs. Because both participants had no problem dealing with the more common C syntax and concepts, the software of the platform could be improved by restricting the top layer to common C syntax only.

Both Sofia and Ron indicated that, of all information that they were provided with, they used the example code in the manual most; Sofia indicated that she would typically start from the example code, because that way she was sure she had the syntax right, and would expand the code from there.

We observed that the basic functionality offered in the software of each component of the platform was sufficient for the participants to get started. However, we found that they would have benefitted if the software library had offered more game-specific code as well. For example, the code recurring most was functionality to automatically identify, at the start of the game, which devices were present for that game. Though both participants found ways to deal with this, it cost them a considerable amount of implementation time. This kind of functionality is common for more games, and in the future the platform software could be extended to offer this functionality as well.

Designing the game software and managing complexity

More complex interactions, for example using the radio to setup a 'dialog' between two devices, took some tries to get right. Both Ron and Sofia especially seemed to struggle getting the structure (the architecture) of their code right, as neither has received formal training in designing the software architecture. As a consequence, they did not consciously think of designing the structure of their code beforehand, sometimes resulting in unnecessarily complex code. Similar difficulties that end-user programmers have with design of their software have also been observed by Myers et a. (2006) and Ko et al. (2011). Typically, both participants would start with a small piece of code, and would keep expanding that code, often resulting in unnecessarily complex (and sometimes unmanageable) code. For example, we observed the code in Sofia's game grew more complex because more and more conditions were added. At some point the code became very hard to understand, as it contained so many if-then-else loops. At that point we introduced to Sofia the concept of a state machine*, and we demonstrated implementing a simple one in her code, by creating functions that each mapped to a state. This immediately improved the structure of her code, and once shown how, she was able to extend the state machine for new states. Of course, it is impossible to expect from non-trained programmers to have knowledge of software design patterns or software behavior abstractions like state machines. However, since both participants seemed to rely heavily on example code, it appears useful to enrich the manual with relevant examples of simple implementations of design patterns. This way, they could program these design patterns by example.

* http://en.wikipedia.org/wiki/Finite-state_machine

We have observed that though the participants had a clear idea of what they wanted to accomplish in their game, they sometimes had trouble finding a way to efficiently implement it. For example, for his application, Ron wanted two devices to engage in a 'dialog'. This meant that device A would play a sound file, while device B would 'listen', and then vice versa. Ron's approach was to have A send a message to B indicating which sound file A was playing. So, B could look up the length of the particular sound file, to find out how long he should be 'listening' before answering. Ron decided to add a look-up list in his code with the length of each sound file. This approach however was possibly buggy; as mistakes were easily made e.g., by forgetting to update length information when new sound files are added to the system. After discussing with Ron we proposed a more simple and robust solution: instead of basing the behavior of the devices on the length of each sound file, device A could simply send a message to device B at the end of his 'talk'. This way, device B merely needs to wait for that signal before responding. We observed that issues, such as the one described, mostly arose when participants were dealing with communications in their games. When two (or more) devices are involved, the designer is forced to consider the result of his code from different 'perspectives' of each device and to consider timing and turn taking issues. It would be an interesting development for the platform to classify recurring patterns and provide ready made solutions in the API thus reducing complexity for designers.

Supporting debugging and testing the prototype

One of the barriers for end-user programmers that Ko et al. (2004) identify is what they call the understanding barrier. Or, evaluating the program's behavior, including compile and run-time errors. We observed similar difficulties in our study: for Sofia most of the compiler errors were insurmountable; she did not understand the errors the compiler gave and was unable to fix the errors. More difficult were the compiler warnings. Though the Eclipse IDE did show these, Sofia mostly ignored them. The Arduino IDE does not even display compiler warnings, resulting in Ron's code breaking down in run-time as the compiler automatically type-casted a `long` in an `int` without warning, losing half of his information. He needed our help to track down this issue, as the Arduino IDE did not provide any feedback at all. The first solution that comes to mind to avoid such problems as described above is to have the IDE perform stricter type checking at compile time. Alternatively, for ease of use and a less steep learning curve for designers, a programming language could be considered that is type-free, or one that dynamically infers types at run-time (Tratt, 2009). However, this approach is not necessarily the ideal solution. Such languages are not yet very common for embedded systems and can introduce erroneous programming behavior when the programmer is not fully aware of the type casting logic implemented by the compiler.

Even more complex for our participants were run-time issues, i.e. finding out why the devices did not respond as they had thought they would. For end-user programmers in general it is a challenge to debug their code in a dedicated environment, in our case it was virtually impossible, as the code runs on embedded hardware. Mostly, debugging for our participants consisted of adding many print statements to the serial port, which at run-time would show on a serial port reader on the computer. Also, a complicating factor was the fact that there are many sources an error can originate from: it could be a mistake in the game logic (code never reaching a certain line), a programming error (e.g., the aforementioned `long` to `int` type-cast) or a failure in the hardware. We observed that Sofia did not follow a systematic approach for tracking down errors; particularly Sofia had trouble narrowing down the lines of code that were generating the errors. Ron had less trouble, he mostly started with testing small pieces of code, and once they had proven to work, he would add them to the rest of the code.

Though not directly related to the hardware or software of the platform itself, an obvious improvement of the use of the platform, would be a tool that better supports debugging run-time information.

PARTICIPANTS' REACTION TO OUR FINDINGS

It is considered good practice in an informative case study setup to discuss the findings with the participants (Klein and Myers, 1999). Sofia agreed to all of the conclusions as reported in this chapter. Ron had a few small remarks, concerning some of the conclusions we drew based on our observations. After discussing with him we agreed with his remarks and changed the text accordingly.

REFLECTING ON THE EVALUATION SETUP

Using the concept maps method as proposed by Gerken et al. (2011) has proven to be a valuable way to obtain insight on many aspects of the uptake of the platform by our participants. We did adapt the method slightly: Gerken et al. let their participants extend and modify the same map every session. We decided not to do this; we tried it once, but the map became too messy. Instead, we let participants create a new map each time, while giving them access to their map of the previous session. We argue that this gives a clearer picture on what the participants have focused on during the last few days of implementation. For example, in Ron's map, if viewed over time, we see the platform components "disappear" over time. They were still in his code, but not so much important for him anymore as that part of the code was working for him already; for us this was an indication that the platform was supporting Ron in his design process. Further, avoiding re-use of the earlier map avoids the tendency of participants to attempt to be consistent with earlier responses. We would rather they give us their mental model at the point in time that they are interviewed rather than they try to 'salvage' the earlier concept map.

Gerken et al. (2011) used graph analysis tools to digitize and analyze the concept maps, e.g., for visualizing changes over time. We did not employ such tools, mostly because in our case the amount of data was much smaller than in the study of Gerken et al. and we were able to simply go over the limited set op concept maps and spot differences over time manually. Furthermore, Gerken et al. observed that the maps could also be used as "helpful prompt" during the interviews and we agree to this observation. We indeed found the comments elicited while making the map very useful for gathering detailed insight and argue that we would not have obtained the same level of detail in an interview without the maps.

By asking the participants to create a concept map and revisit their pitch every session, we unavoidably influenced the design process. Gerken et al. (2011) have made a similar observation. Both participants reflected on this, saying that the concept maps and pitches helped them to "take a step back" and see the bigger picture; something they would have been less inclined to do in their usual design processes. The question of course is whether it has influenced the process to such a degree that the conclusions that we draw should be attributed to the map making activity alone. We argue that its influence was limited: our main conclusions on the platform influencing the design process related to the speed with which the participants were able to create working prototypes and to the fact that the platform allows participants to freely explore the technology. Arguably, the map making activity has little effect on either of these observations and thus can be disregarded.

At the start of the study, we limited our role to technical support only, as to not influence the design process that we were studying. At some points during the study this turned out to be difficult. For example, it occurred that one of the participants wanted to implement a certain feature in the game in a way that we, from experience, knew would be difficult for children to understand. For some of these issues we decided to step out of our self-assigned roles and give feedback on the game design. Had we set up our study with pre-set tasks, instead of involving participants who were working on their final Master's project, we would not have done so. However, in our case both participants' end result was of significant importance to them, and we found it nearly un-ethical to withhold our expert opinion when that could help them to significantly improve the chances of success of their game designs.

Finally, we identified the threat to this study that users might answer favorably only, to please us. We have made an effort to counteract this, e.g., by encouraging and discussing openly their critical remarks, and we can conclude that we have not observed such instances. This might be partly due to the fact that participants were working on their own projects, and were not working on pre-set tasks. This way, it was in their best interest to also report negative experiences, as they could use our help to advance past these issues to successfully finish their projects.

CONCLUSION

In this paper we have presented RaPIDO, a rapid prototyping tool for prototyping physical interaction for outdoor games. RaPIDO was designed to address a set of requirements from prototyping technology that we have identified in our own experiences with designing outdoor games for children, but also that are derived more generally by considering the challenges facing designers in this field.

To verify our assumption that RaPIDO could support designers in their design process, we executed an interpretive case study in which we followed two designers. Our expectations were largely met; we concluded that indeed designers benefitted from a platform like RaPIDO as it allowed them to quickly create interactive prototypes, more quickly than they would have if they had had to put together the hardware themselves. Furthermore, it allowed the designers to freely explore the technology, developing the game design and interactive technology concurrently, which we argue is beneficial for the resulting game experience. The technology provided by the platform was found sufficient to support the games the designers aimed to create. Most issues identified in this study related to the software of the platform. The designers were able to create interactive games, though they ran into problems once the complexity of their code increased.

Reflecting on the design decisions we made for RaPIDO, more specifically our decision to leave out GPS and a display, we can say the following: for the games we have created earlier we did not need GPS, and we made the same observation in the case study; the designers did not need it in their games. Regarding leaving out a display we note that it, too, was not necessary in the games created. However, we did observe that for debugging and testing purposes, a small display on each device that would display some information about the current state of the device could have been helpful during testing. Alternatively, an external interface, e.g., on a tablet, could be designed to review the internal state of the devices.

As the HCI field is broadening its scope from a largely desktop-centered context to also include tangible, embodied and ubiquitous contexts, we argue there is a need for tools like RaPIDO. Portable technology and embedded hardware are becoming more main stream and tools that can support designers to rapidly prototype (embedded) interactions are therefore becoming increasingly important. To draw a parallel: in the field of desktop interaction currently many tools are available for interaction designers to quickly sketch interactive screen applications (e.g., Balsamiq*), allowing interaction designers to focus on designing the interaction, while not getting distracted by technical implementation details. Tools like Arduino offer designers a means to create interactive physical prototypes, however, to take full advantage of what embedded technology can offer nowadays, is often beyond the skills of interaction designers. As such, the creation of RaPIDO is an added asset to the tools that designers can use to easily create fully functioning prototypes. In contrast to other rapid prototyping tools, RaPIDO does not offer separate components that can be assembled by a designer to create a prototype. Instead all components are integrated on a dedicated, compact PCB creating a small and portable solution. Furthermore, a software API is offered for programming the devices,

* http://www.balsamiq.com/

providing an abstract layer between the actual hardware and the application software, so designers can focus on programming the game interactions.

Though our initial motivation to develop RaPIDO was to support rapid development of outdoor games for children, we are confident that the platform has a wider scope than that: we expect that the platform can also be used to prototype physical interactions that do not rely on screen based interaction and access to a computer network. It could very well be used as an exploration tool for designers to start examining interactions and thereby hands-on exploring the design space, without the need to construct technology first.

As we have seen in Chapter 3 and also in the user study presented in this chapter, the design process benefits from a platform that can support the creation of robust, functioning prototypes in the early stages of a design process. Regarding the more general field of pervasive and ubiquitous computing, Tang et al. (2011) arrive at a similar conclusion: "In pervasive computing, many complicated and unpredictable interactions exist between system components and users, it would be, thus, best to start with something richly interactive". However, Carter et al (2008) conclude from a literature survey and interviews with developers that rarely any iterations in the design process are performed and that this is a threat to the ecological validity of evaluations done with ubicomp systems. One of the main challenges they identify is the creation of robust and reliable prototypes. Though high-functioning prototypes are indeed created, these are often fragile and can only be evaluated in a safe lab-environment. Apparently, the current offering of prototyping platforms does not yet enable developers and designers to create sufficiently robust prototypes.

Therefore, if designers want to rapidly explore new interaction possibilities and at the same time be able to evaluate them with users in a real context, new, integrated, prototyping tools like RaPIDO should be developed, creating ready to use platforms for interaction designers to pick up and deploy.

FUTURE WORK

Most effort in this project was spent on creating the embedded hardware, it was a deliberate decision to not invest time in developing a dedicated programming environment, because such a development would have been impossible to complete within the scope of the project. Instead, we chose to offer a software API in C, as we expected our target users to have some programming skills.

We concluded from our study that designers indeed were able to appropriate the software libraries and effectively use the platform. However, as they had not followed formal training in software engineering, it is not a surprise that many of the issues we found related to creating more complex programs. For example, conceiving a state-machine was beyond the abilities of our participants.

To address this, effort should be spent to close the gap between the level of abstraction offered by the software and the level of complexity that designers can be expected to handle. First, the API could be extended to offer more game-related functionality. Though some game-generic functionality is already offered, in the study we observed that parts of the code written by the participants were generic for Head Up Games (e.g., detecting what devices are present at the start of a game). Including such game-generic functionality in the API could very well reduce the complexity for the designers. Furthermore, the Head Up Games designed so far seem to be well suited to be represented as state-machines. The current API could be extended to offer a basic structure of a state-machine to facilitate this.

For the type of designers we were aiming at, we argue that a textual language offers them the most flexibility for creating their games. However, to further accelerate the creation of prototypes by designers, or alternatively, if we look beyond designers, and consider for example parents, schoolteachers or even children themselves as game creators, a "sketch"-like, visual interface could be explored to support the creation process. An interesting example of a visual programming environment for Arduinos, though still in development, is ModKit*. ModKit is a spin-off of Scratch (Maloney et al., 2010), a visual programming language especially designed for children between 8 and 16. ModKit offers a drag-and-drop interface to assemble a program quickly, which considerably reduces the chances of syntax errors. However, ModKit is targeted at users who also create the electronics of their projects and it offers

* http://www.modk.it/

generic building blocks to create software structures. In our case the hardware is rather fixed and we aim to offer dedicated functions that support rapid creation of (gaming) interactions. Still, it would be interesting to see if a tool as ModKit can be extended to offer dedicated game-related building blocks and whether or not that would better support designers to create their interactive prototypes.

During the development of the prototypes we have learned many valuable lessons about the implementation of specific hardware components. Also, in our attempts to keep costs down, RaPIDO turned out to have a larger footprint than we initially desired. Our aim is to create a second iteration of the hardware and simultaneously reduce the size of the devices, making them more portable and appropriate for a mobile context.

REFERENCES

Avrahami, D., and Hudson, S. E. (2002). Forming interactivity: a tool for rapid prototyping of physical interactive products. In Proc. of the 4th Conf. on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (pp. 141–146).

Ballagas, R., Ringel, M., Stone, M., and Borchers, J. (2003). iStuff: a physical user interface toolkit for ubiquitous computing environments. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 537–544).

Bekker, T., and Sturm, J. (2009). Stimulating physical and social activity through open-ended play. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 309–312).

Bekker, T., Sturm, J., and Eggen, B. (2010). Designing playful interactions for social interaction and physical play. Personal and Ubiquitous Computing, 14(5), 385–396.

Benford, S., Rowland, D., Flintham, M., Drozd, A., Hull, R., Reid, J., Morrison, J., and Facer, K. (2005). Life on the edge: supporting collaboration in location-based experiences. In Proc. of CHI ’05 (pp. 721–730). ACM, Portland, Oregon, USA.

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

Buxton, B. (2007). Sketching User Experiences: Getting the Design Right and the Right Design (1st ed.). Morgan Kaufmann.

Cao, J., Riche, Y., Wiedenbeck, S., Burnett, M., and Grigoreanu, V. (2010). End-user mashup programming: through the design lens. In Proc. of the 28th Int. Conf. on Human Factors in Computing Systems (pp. 1009–1018).

Carter, S., Mankoff, J., Klemmer, S. R., & Matthews, T. (2008). Exiting the Cleanroom: On Ecological Validity and Ubiquitous Computing. Human–Computer Interaction, 23(1), 47–99.

Costikyan, G. (2002). I have no words & I must design: Toward a critical vocabulary for games. In Computer Games and Digital Cultures Conf. Proc. (pp. 9–33).

Dahlbäck, N., Jönsson, A., and Ahrenberg, L. (1993). Wizard of Oz studies: why and how. In Proc. of the 1st Int. Conf. on Intelligent User Interfaces (pp. 193–200). .

Engebretson, A., and Wiedenbeck, S. (2002). Novice comprehension of programs using task-specific and non-task-specific constructs (pp. 11–18).

Gerken, J., Jetter, H.-C., Zöllner, M., Mader, M., and Reiterer, H. (2011). The concept maps method as a tool to evaluate the usability of APIs. In Proc. of the 2011 Annual Conf. on Human Factors in Computing Systems (pp. 3373–3382).

Greenberg, S., and Fitchett, C. (2001). Phidgets: easy development of physical interfaces through physical widgets. In Proc. of the 14th Annual ACM Symposium on User Interface Software and Technology (pp. 209–218). ACM, New York, NY, USA.

Hendrix, K., Yang, G., Van de Mortel, D., Tijs, T., and Markopoulos, P. (2008). Designing a Head-Up Game for Children. In Proc. of the HCI'08 Conf. on People and Computers XXII (Vol. 1, pp. 45–53).

Hodges, S., Villar, N., Scott, J., & Schmidt, A. (2012). A New Era for Ubicomp Development. IEEE Pervasive Computing, 11(1), 5–9. doi:10.1109/MPRV.2012.1

Klein, H. K., and Myers, M. D. (1999). A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. MIS Quarterly, 23(1), 67–93.

Ko, A.J., Myers, B. A., and Aung, H. H. (2004). Six Learning Barriers in End-User Programming Systems. In Visual Languages and Human Centric Computing, 2004 IEEE Symposium On (pp. 199 –206).

Ko, Andrew J., Abraham, R., Beckwith, L., Blackwell, A., Burnett, M., Erwig, M., Scaffidi, C., Lawrance, J., Lieberman, H., Myers, B., Rosson, M. B., Rothermel, G., Shaw, M., and Wiedenbeck, S. (2011). The state of the art in end-user software engineering. ACM Comput. Surv., 43(3), 21:1–21:44.

Landay, J. A., & Myers, B. A. (2001). Sketching interfaces: toward more human interface design. Computer, 34(3), 56–64. doi:10.1109/2.910894

Lee, J. C., Avrahami, D., Hudson, S. E., Forlizzi, J., Dietz, P. H., and Leigh, D. (2004). The calder toolkit: wired and wireless components for rapidly prototyping interactive devices. In Proc. of the 5th Conf. on Designing Interactive Systems (pp. 167–175).

Li, Y., Hong, J. I., & Landay, J. A. (2004). Topiary: a tool for prototyping location-enhanced applications. In Proc. of the 17th annual ACM sym. on User interface software and technology (pp. 217–226).

Lindley, S. E., Le Couteur, J., and Berthouze, N. L. (2008). Stirring up experience through movement in game play: effects on engagement and social behaviour. In Proc. of the 26th Annual SIGCHI Conf. on Human Factors in Computing Systems (pp. 511–514).

Magerkurth, C., Cheok, A. D., Mandryk, R. L., and Nilsen, T. (2005). Pervasive games: bringing computer entertainment back to the real world. Comput. Entertain., 3(3), 4–4.

Magerkurth, C., Engelke, T., and Grollman, D. (2006). A component based architecture for distributed, pervasive gaming applications. In Proc. of the 2006 ACM SIGCHI Int. Conf. on Advances in Computer Entertainment Technology.

Magielse, R., and Markopoulos, P. (2009). *HeartBeat*: an outdoor pervasive game for children. In Proc. of the 27th Int. Conf. on Human Factors in Computing Systems (pp. 2181–2184).

Maloney, J., Resnick, M., Rusk, N., Silverman, B., and Eastmond, E. (2010). The Scratch Programming Language and Environment. Trans. Comput. Educ., 10(4), 16:1–16:15.

Mueller, F., Agamanolis, S., and Picard, R. (2003). Exertion interfaces: sports over a distance for social bonding and fun. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 561–568). ACM, Ft. Lauderdale, Florida, USA.

Mueller, F., Stevens, G., Thorogood, A., O'Brien, S., and Wulf, V. (2007). Sports over a Distance. Personal and Ubiquitous Computing, 11(8), 633–645.

Myers, B. A., Ko, A. J., and Burnett, M. M. (2006). Invited research overview: end-user programming. In CHI '06 Extended Abstracts on Human Factors in Computing Systems (pp. 75–80).

Obrenovic, Zeljko, & Martens, J.-B. (2011). Sketching interactive systems with sketchify. ACM Trans. Comput.-Hum. Interact., 18(1), 4:1–4:38. doi:10.1145/1959022.1959026

Quevedo-Fernández, J., & Martens, J. B. O. . (2012). Demonstrating idAnimate: a multi-touch system for sketching and rapidly manipulating animations. In Proc. of the 7th Nordic Conf. on Human-Computer Interaction: Making Sense Through Design (pp. 767–768). New York, NY, USA: ACM.

Raychoudhury, V., Cao J., Kumar, M., Zhang, D. (2012) Middleware for pervasive computing: A survey, Pervasive and Mobile Computing, doi:10.1016/j.pmcj.2012.08.006

Rennick Egglestone, S., Walker, B., Marshall, J., Benford, S., and McAuley, D. (2011). Analysing the Playground: Sensitizing Concepts to Inform Systems That Promote Playful Interaction. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler (Eds.), INTERACT 2011 (Vol. 6946, pp. 452–469).

Salen, K., and Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals (illustrated edition.). The MIT Press.

Scaife, M., Rogers, Y., Aldrich, F., and Davies, M. (1997). Designing for or designing with? Informant design for interactive learning environments. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 343–350).

Soler-Adillon, J., and Parés, N. (2009). Interactive slide: an interactive playground to promote physical activity and socialization of children. In Proc. of the 27th Int. Conf. Extended Abstracts on Human Factors in Computing Systems (pp. 2407–2416).

Soute, I., Kaptein, M., and Markopoulos, P. (2009). Evaluating outdoor play for children: virtual vs. tangible game objects in pervasive games. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 250–253). ACM, Como, Italy.

Soute, I., Lagerstrom, S., and Markopoulos, P. (2013). Rapid Prototyping of Outdoor Games for Children in an Iterative Design Process.

Soute, I., Markopoulos, P., and Magielse, R. (2009). Head Up Games: combining the best of both worlds by merging traditional and digital play. Personal and Ubiquitous Computing, 14(5), 435–444.

Tang, L., Yu, Z., Zhou, X., Wang, H., & Becker, C. (2011). Supporting rapid design and evaluation of pervasive applications: challenges and solutions. Personal and Ubiquitous Computing, 15(3), 253–269. doi:10.1007/s00779-010-0332-6

Tratt, L. (2009). Dynamically Typed Languages. (M. V. Zelkowitz, Ed.)Advances in Computers, 77, 149–184.

Valk, L. de, Rijnbout, P., Bekker, T., Eggen, B., Graaf, M. de, and Schouten, B. (2012). Designing for playful experiences in open-ended intelligent play environments. In Proc. of IADIS Int. Conf. Game and Entertainment Technologies (pp. 3–10).

Verhaegh, J., Soute, I., Kessels, A., and Markopoulos, P. (2006). On the design of *Camelot*, an outdoor game for children. In Proc. of the 2006 Conf. on Interaction Design and Children (pp. 9–16).

Villar, N., Gilleade, K. M., Ramdunyellis, D., and Gellersen, H. (2007). The VoodooIO gaming kit: a real-time adaptable gaming controller. Comput. Entertain., 5(3).

Virzi, R. A. (1989). What can you learn from a low-fidelity prototype? In Proc. of the Human Factors and Ergonomics Society Annual Meeting (Vol. 33, pp. 224–228).

Chapter 5

Evaluating player experience in Head Up Games

This chapter is based on:

Soute, I., Bakker, S., Magielse, R., and Markopoulos, P. (2013). Evaluating player experience for children's outdoor pervasive games. Entertainment Computing, 4(1), 25–38.

ABSTRACT

There is a growing body of research in pervasive outdoor gaming, mainly focused on adult players playing games on smart phones. Published evaluations of the player experience in such games are largely based on anecdotal descriptions and post-play surveys. The latter approach is especially challenging to apply when the playtest participants are children. Observations of game play so far have been ad hoc relying on unstructured observation, which makes it difficult to extract reliable conclusions from observations and to draw comparisons between different games. In this paper we discuss two methods developed specifically for evaluating the player experience in children's outdoor games: the Outdoor Play Observation Scheme (OPOS) by Bakker et al. (2008) and GroupSorter. We discuss their application in three case studies and conclude that OPOS is useful in quantifying the different types of play behavior in outdoor games; GroupSorter generates qualitative data on the play experience. Moreover, the application of GroupSorter is not limited to game development but can be used for obtaining user input in other context as well.

INTRODUCTION

A consequence of the current wide adoption of mobile computing is the emergence of mobile and pervasive gaming, where mobile interactive devices with computing and communication capabilities (typically smart phones) are used to play games outdoors. These games may involve one or more players who may be distributed or co-located, and where game play can take place in the broad variety of locations and contexts where one might expect such mobile devices to be used. This class of games represents a vast new growth area for mobile interactive technologies but also, we argue, a new set of methodological challenges relating to the user centered design of related systems. Research in this area has progressed largely in terms of developing research prototypes and charting the related technical, interaction, and game design challenges. For example, the pioneering pervasive game *Can You See Me Now*? (Benford et al., 2006) is a mixed reality chasing game intended for adults dispersed in an urban environment; the emphasis of the researchers was on exploring and demonstrating the limitations but also the opportunities related to creating pervasive games that rely on Wi-Fi and GPS infrastructures. Similar examples are CatchBob! (Nova et al., 2006) and Feeding Yoshi (Bell et al., 2006). In these projects the evaluations are focused more on the technological innovations, and on exploring the nature of the emerging user experiences and less on the ambition to just create a fun and playable game.

As the transition is made between the initial pioneering phase to more routine development and eventually adoption of such games, the need emerges for a suitable evaluation methodology. To this point there has been no systematic effort on this front. Partly, this is a consequence of the novelty of the field: by its nature methodology research inevitably lags behind design innovations. Then again, it could reflect expectations by designers and researchers, that traditional user centered design methodology (Norman and Draper, 1986) or more recently experience design (Hassenzahl, 2011) suffices for the purposes of mobile and pervasive game design, perhaps with minor or major adaptations to fit the specific design context. While for many cases this may hold true, our particular interest in Head Up Games leads us to a different position.

Similar to the motivations of Head Up Games, other researchers have explored interactive games and playful installations, where children can play together and where physical activity is an inherent part of the play experience. Examples of such research are interactive games that support exertion (Mueller et al., 2003), interactive installations that support groups to by physically active such as the interactive slide or the interactive fountain (Soler-Adillon and Parés, 2009), the enhancing sporting with sensor based interactive technology (Bekker and Eggen, 2008), etc. A common goal for all these genres of interactive technology is physical activity and social interaction and, of course, the designer’s intent to embed these in an engaging and fun activity. These shared design goals bring about recurring methodological challenges with regards to evaluation. The involvement of children already presents a distinct set of methodological challenges for evaluators that require either the adaptation of evaluation methodologies originally developed for adult test participants or the invention of novel evaluation methods (Markopoulos et al., 2008). Further, the mobile context, the physical activity and the open space, limit the applicability of testing methods originally intended for a laboratory context and a static interaction platform such as a desktop computer.

In the next section we elaborate these two points, in order to motivate the development of specialized methodologies for testing mobile games for children and we review relevant, existing research in similar areas. Next, we describe the Outdoor Play Observation Scheme (OPOS), a structured observation method for evaluating emerging play behaviors. Furthermore, we present GroupSorter, a method we have developed specifically for evaluating the experienced fun as well as the rationale behind it. GroupSorter is suitable for children age 7 and up. Then, we provide details on the application of both methods in three case studies. Subsequently, we reflect on the usage of the methods and provide ideas for further improvement.

RELATED WORK

This section is divided in three parts: a review of existing evaluation methods for play, games and physical activity in general, a review of games that have been developed that are closely related to Head Up Games, and, concluding, a review of child-centered usability evaluation methods.

RELATED EVALUATION METHODS

Let us start this review by stating that there are many different definitions of both 'play' and 'games' and that there exists a 'surprisingly complex relationship' (Salen and Zimmerman, 2003, p. 72) between the two. Furthermore, there are many theories on why play exists and the role of play in child development; for an overview of both classic as well as modern theories we refer the reader to (Mellou, 1994). As there is no consensus on the exact definition of 'play' it is not hard to see there exists no 'golden standard' for assessing play, and, unsurprisingly, many different approaches in as many different disciplines are taken. We highlight several of these, though this is by no means an exhaustive review.

Study of play

In the area of developmental psychology play is regarded as an essential aspect in a child's development (Scarlett et al., 2004) and studies of children's play often focus on the social development of young children. Assessment of children's play is most commonly done by observation.

For instance, Parten (1932) conducted one of the very first observational studies of children's play. She defined an observation scheme to study the social participation of preschoolers (2-4yr) in spontaneous play using a one-minute sampling method. For four months, children were observed one hour per day in which they were free to choose what and with whom to play. Parten defined a scale for classifying children's social participation in play, ranging from non-social play ('solitary play') to play involving high social participation ('cooperative or organized supplementary play'). From her observations she concluded that the social participation increases with age (Parten, 1932).

Rubin's Play Observation Scheme (POS) (Rubin, 2001) combined Parten's work with the Smilansky classification of play behavior in an observation scheme that classifies both social play and cognitive play. POS has been used in several projects investigating the free play behavior of preschoolers, e.g., Rubin et al. (1978) and Hetherington et al. (1979). Rubin applied POS to compare free-play behaviors of preschool- and kindergarten-aged children. One of the main conclusions that was drawn, was that kindergarten children engage more in group and dramatic play than preschool children. Hetherington et al. (1979) applied a slightly modified version of POS to study the effect of divorce on social interaction and play on preschool children. The children were observed during 6 sessions over the course of 2 years after their parents divorced. Based on the results Hetherington et al. concluded, amongst others, that compared to children of non-divorced families play patterns of children from divorced families were more fragmented and less socially mature during the first year after the divorce.

Metin (2003) studied children's play in a playground and the effect of the equipment design on their play behavior. She classified behavior in behavior patterns (e.g., talking, pretending) and play types (e.g., sensorimotor play, pretend play, games with rules). 70 children, aged 6 to 12, were observed in a park; after the observation session a short interview was conducted. The results of the study showed that today's playground has little value in terms of play. Children's physical and social developments are supported to an extent, however cognitive and emotional development are not fostered.

Study of games

Up until the arrival of computer games, the study of games was sparse. A few notable exceptions are Johan Huizinga (1955) and Roger Caillois (2001) who discussed games from the perspective of culture and sociology. However, as computer games rose in popularity, the interest from the academic field into gaming also grew. For example, the influence of violent video games on aggression has often been studied (Anderson and Carnagey, 2009; Sherry, 2001). Experiments typically use questionnaire data, sometimes combined with physiological measurements, to assess the participant's attitude towards violence, aggressive behavior, aggressive affect or aggressive cognitions.

The application of digital games for children has often been studied from an educational perspective (Kirriemuir and McFarlane, 2004). Typically in the field of HCI educational games are evaluated using cognitive tests, questionnaires (e.g., Verhaegh et al. (2012)), and interviews (e.g., Parkes et al. (2008)).

Social aspects of video games that are studied are for example the feeling of presence (De Kort et al., 2007), player enjoyment and engagement (Chen et al., 2006). De Kort et al. developed the Social Presence in Gaming questionnaire (SPGQ) to measure the player's feeling of social presence in digital games. Chen et al. interviewed MMORG (Massively Mulitplayer Online Role-playing Games) in a semi-structured interview to get a holistic account of the players gaming experience.

Finally, digital games are also researched to find out what makes them fun for children (Barendregt et al., 2007). To this end Barendregt et al. have developed the PIPC method: the problem identification picture cards method. This method combines the traditional think aloud method with picture cards that indicate certain types of usability or fun problems. Children can place the appropriate card in a box when they encounter a problem while playing a computer game.

Physical activity in a play context

As it is becoming more and more clear that a sedentary lifestyle can cause health problems, research into encouraging physical activity is growing. Dollman et al. (2009) list eight approaches for physical activity assessment of children, such as heart rate monitoring, accelerometry and direct observation. Each of the tools has its own merits and limits with respect to both validity of the measures as well as the practical applicability. Mostly, these tools are applied in the context of obesity research, and in clinical applications.

More specifically in the context of outdoor play, Haug et al. (2010) use questionnaires administered to principals and students in order to determine the relationship between the outdoor environment and the participation in physical activity during school breaks of children age 8-15. Furthermore, the SOPLAY observation scheme (McKenzie et al., 2000) records play activities of groups. Every interval, typically every few minutes, an observer scans a target area, noting the number of participants, their physical activity level and other contextual characteristics.

Conclusion

Summarizing we can say that play, games and physical activity are studied in many different domains from many different perspectives. Though a variety of methods are applied, we conclude that observations, questionnaires, and interviews are most commonly used.

RELATED EVALUATION STUDIES

In this section we will look closer at a selection of games that have a close relation to Head Up Games, in particular with respect to the evaluation methods used in these projects.

The field of pervasive games is growing rapidly. Examples of pervasive games for adults are *Can You See Me Now*? (Benford et al., 2006) and *PacManhattan* (Lantz, 2007). Both games are location-based games, targeted at adults. Both games mix online and offline play to create a chase game that is played out in an actual cityscape. *Can You See Me Now*? was evaluated using naturalistic observations, by reviewing system logs of usage and through discussions with participants. For *PacManhattan* it is unclear how it was evaluated.

Pervasive games that have been specifically designed for children are *Savannah* (Benford et al., 2005) and the *Hunting of the Snark* (Rogers et al., 2002). Both games are collaborative educational games for children. The evaluation of *Savannah* entailed unstructured video observation of the game play combined with system logs of the game interactions. Based on these data usability issues with the system were identified. In the *Hunting of the Snark* a more structured observation scheme was used that focused on various aspects of playful learning.

An example of an *interactive play objects* (Bekker and Eggen, 2008) are the *ColorFlares* (Bekker and Sturm, 2009). A user test was carried out with *ColorFlares* to test whether children created their own games and to see whether open-ended play stimulated social interaction. Play sessions were recorded for analysis of play behavior i.e. the type and number of games children created. After the play session children were asked to fill in a questionnaire. The study showed that children were in fact able to come up with diverse games that incorporated the *ColorFlares*, and that the children enjoyed doing this.

The game *Breakout for Two* (Mueller et al., 2003), an exertion interface, is a sports game that is played by two players connected through the Internet. The game has been evaluated using questionnaires and interviews. Another example of a game that requires more physical activity is the adaptation of the Nintendo game *Donkey Konga* (Lindley et al., 2008), where bongos as input devices replaced the standard game controllers. This game was evaluated in a study comparing the standard game controllers with the bongos. The data was gathered using questionnaires to measure engagement and by analyzing video footage to measure the amount of gestures and social interaction. To analyze the social interaction definitions based on the Autism Diagnostic Observation Schedule were used. Based on the analysis of the video footage and the questionnaires the researchers concluded that the social interaction in the game with the bongos was significantly higher, and that this did not detract from the engagement in the game.

RELATED METHODS IN USABILITY EVALUATIONS WITH CHILDREN

In this section we review methods for usability evaluation with children and discuss their applicability in outdoor games.

Markopoulos et al. (2008) divide usability evaluation methods for and with children roughly into: observational methods, verbalization methods, survey methods, and inspection methods.

Verbalization methods

Evaluation of interactive systems with verbalization methods such as the Think-Aloud method (Boren, 1999) aim to capture the thought processes of test-participants, by asking them to verbalize these during a test session. Think-aloud is often used in usability testing, and research has shown that it is an effective means to identify improvements for interaction design (McDonald et al., 2012). Nevertheless it has some well-known drawbacks. The first drawback concerns the extra cognitive load required for test participants to verbalize their thoughts. The second drawback is regarding the difficult social situations that might arise when a test-facilitator attempts to avoid interacting with test participants to avoid influencing their thought processes (Boren, 1999). These two core issues are even more pronounced when test participants are children, raising doubts regarding the applicability and utility of the method for children. Research has established that it is feasible and effective also with children participants above age 9 (Donker and Markopoulos, 2002), though this method assumes that the participants can talk while interacting with the product, can spare the cognitive resources and can be heard talking while they do so, which is possible when they remain stationary in one location during the course of the evaluation. These conditions are less likely to hold during the evaluation of a Head Up Game, which are often medium to high paced activities and are played with many players. Furthermore, the game play makes it virtually impossible to verbalize thoughts, as the game itself requires the players to communicate with other players. Consequently, verbalization methods are deemed inappropriate for the type of games that we want to evaluate.

Inspection methods

Inspection methods are methods in which one or more experts in the field analyze the (designed) interaction. In a study on the usefulness of playability heuristics in pervasive game evaluations Jegers (2008) found that although many issues could be identified by heuristics, some major issues were not identified. For similar reasons, using inspection methods is not common practice in game design. Salen and

Zimmerman (2003, p. 12) state that it is nearly impossible for "even a veteran designer [to] exactly predict what will and will not work before experiencing the game firsthand". Salen and Zimmerman, as well as Fullerton et al. (2004), emphasize the importance of playtesting games. Based on our experiences with game design and evaluation, we agree on this, and therefore conclude that inspection methods are not appropriate for our purpose.

Observation methods

Numerous observation methods for adult test participants are used for evaluation ranging in their rigor and their completeness. There is a focus primarily on identifying usability problems rather than on evaluating fun and the overall experience. In most cases, observation is unstructured although in some cases structured scoring sheets are created to help code observations in a structured way. An example is the DEVAN scheme for coding observations during usability testing with adults (Vermeeren et al., 2002). Such observation schemes that require comparing interaction at a micro level (clicks, selections, etc.), to verbalization and screen contents are hardly able to cope with the fast pace of the outdoor game, the wide range of the playing field, and the intense experience that needs to be evaluated.

Focus on children as test-participants prompted Barendregt et al. (2007) to extend the DEVAN observation scheme with elements referring to playing games. However, such coding schemes are very much tied to the problem-solving aspects of learning a new game and thus focus on how displays are interpreted and how the user can figure out how to proceed with interaction; this cognitive element of gaming is still present in outdoor games but less central to the game play and the emerging play experience. Instead, the elements of physical activity and social interaction contribute significantly to the emerging game experience, and therefore are critical for the evaluator. However, both physical activity and social interaction components of games are currently not addressed in these observation schemes.

Survey methods

Interviews or questionnaires are often administered to obtain children's opinions about a specific product or system. Because they are issued after using the product they avoid the problems of verbalization during use as discussed above, and they can equally well be applied for mobile contexts. Both interviews and questionnaires need to be adapted when test participants are children to address their level of language development, interests, and skills. Methodologists have identified several pitfalls when surveying children's opinions during usability testing (Read, 2008): satisficing which occurs when a child gives a more or less superficial response simply to complete their task (e.g., marking randomly one of the multiple choice options in a questionnaire), or suggestibility, where children's responses reflect their perceived influences from the testing context, e.g., adults, environment, etc.

Applying lengthy questionnaires for children can pose a problem as reading skills and the ability to remain focused on a (long) list of questions is not yet properly developed in children. Also, writing skills may vary largely over age groups and so questionnaires for children require special consideration to be taken into account concerning, e.g., length and wording (Markopoulos et al., 2008), but also with regards to the validity of the answers given; as noted by Read and MacFarlane (2006) children tend to give maximum scores indiscriminately to rating scales not allowing useful conclusions to be given with respect to how they evaluate their experience with games. A set of instruments and tips for surveying children's opinions has thus been developed, for a review see (Markopoulos et al., 2008) and these are largely applicable in the present context. These instruments are often used in classroom settings or in laboratory situations and are easy to administer as questionnaires or in one-to-one interviews. Adaptations may be needed when working with a group of children in between outdoor play sessions or just after them, because from a practical point of view, it is often impossible to have one-to-one interviews with each child after the play session. It would simply take too much time, or otherwise many evaluators would need to be present.

CONCLUSION

Concluding we can say that observation and survey methods are the two most likely types of methods to gain useful insights into the play behaviors and play experience and, at the same time, are most appropriate for children in an outdoor context. However, the methods in their original form needed to be tailored to our specific needs, which resulted in the two methods we present in this paper: OPOS and GroupSorter.

METHODS

From the above review and discussion it is clear that there are several requirements that our particular problem domain puts upon evaluation methodology. First, we address children, so the method must be appropriate for them. Second, the children play outdoors and are quite animated during play, so the evaluation needs to fit their play activities, and, finally, to address them as a group: evaluation sessions will by the nature of these games be carried out by groups of children. The opinions of children have to be collected efficiently and avoiding problems such as children talking to each other about the evaluation or even the interview, while each is waiting for his/her turn to be interviewed.

In this section we will elaborate on two methods that are suitable for evaluating computer enabled outdoor games for children. The first method is OPOS (Bakker et al., 2008), an observation scheme for quantifying play behavior. The second method we present is GroupSorter. We developed GroupSorter to gather qualitative data on the children's play experience.

OPOS

Head Up Games are intended to stimulate physical activity and social interaction, which are essential aspects of traditional outdoor children's games. Both physical activity and social interaction are possible to observe directly, so the evaluation of these games (and other related genres of games) should not be limited to surveying children's experience and reporting anecdotes of play behaviors only. Rather, designers and researchers need observational methods that evaluate the extent to which these behaviors emerge during game play. While existing evaluations of pervasive games reported in literature often mention that observations were carried out, these seem to be for the most part unstructured observations, and the analysis procedure is frequently undisclosed (e.g., *Savannah* (Benford et al., 2005), and *Ambient Wood* (Rogers et al., 2004)). While unstructured observation can be sufficient in exploratory evaluations, structured observations can produce more informative and reliable results, particularly for research studies aiming to validate a particular design concept or comparative evaluations of different games, or versions of a game.

Unfortunately, neither the literature on pervasive games nor the related discipline of behavioral analysis provides an appropriate coding scheme for our purposes. Thus, a new observation scheme was developed by Bakker et al. (2008), called the Outdoor Play Observation Scheme (OPOS). OPOS enables the collection of quantifiable data on physical activity and social interaction, and thereby allows a comparison to be drawn between different games (e.g., Head Up Games, traditional outdoor games, computer games). The process of developing OPOS is extensively described by Bakker et al. (2008).

As we have seen several different observations schemes are known for assessing play and physical activity. From these, POS (Rubin, 2001) and SOPLAY (McKenzie et al., 2000) are closest to the purpose and content of OPOS. We will shortly describe similarities and differences between these observation schemes.

Table 5.1 shows an overview of the characteristics of each of the observations schemes. SOPLAY focuses on measuring physical activity of large groups of children and its categories for observation are similar to those of OPOS. With SOPLAY every few minutes an observer performs a scan of the play-area and players, and notes what type of activity children are engaged in. Though the categories of physical activity in SOPLAY and OPOS are similar, the unit of analysis and

Table 5.1 Overview of observation schemes

	POS	OPOS	SOPLAY
Aim	Evaluate social play development of child	Evaluate play behavior triggered by computer enabled games	Evaluate physical activity in leisure time
Observation type	Direct, naturalistic observation	Indirect, naturalistic observation	Direct, naturalistic observation
Unit of analysis	One child	Each child in group of children	Group of children
Context	Free play	Outdoor game play	Leisure time in school
Data collection	Time-sampling (partial interval recording)	Duration recording and event recording	Time-sampling (momentary)

the sampling frequency at which the behavior is observed are different. SOPLAY observes the children as a group resulting in an overall figure of how active the group is; in contrast in OPOS we observe each child separately. Also, we collect the data at a higher frequency, i.e. over seconds rather than over minutes. The game play may only last a few minutes in total, rendering a sampling frequency of every few minutes useless. Both collecting the data for each child instead of for the whole group as well as sampling at a higher frequency is needed for the benefit of the designer or evaluator of the game: knowing that the game leads to a walking behavior overall is important, but does not help the designer evaluate specific design decisions or elements in the game.

POS	OPOS	SOPLAY
Social	Physical Activity	Physical activity
Solitary play	Intensive	Sedentary
Parallel play	Non-intensive	Walking
Group play*	No physical activity	Very active
Cognitive	Focus	
Functional play	Looking at other players	
Constructive play	Looking at game objects	
Exploration	Looking at something out of sight	
Dramatic play	Looking at something else	
Games-with-rules*	Social	
Non-play behaviors	Functional	
Unoccupied	Non-functional positive/neutral	
Onlooker	Non-functional negative	
Translation	With non-player	
Active conversation*	Unintended physical contact	
Agression		
Rough-and-tumble		
Hovering		
Anxious behavior		
Uncodable		
Out of room		

Table 5.2 Categories and subcategories of the observation schemes

POS is an extensive observation scheme to evaluate children's play in all of its forms. It encompasses the whole range of play behavior in the context of social development and interaction. In our Head Up Games some of these behaviors are by their nature not present, such as solitary play. Rather, in a typical Head Up Game only the behaviors listed in Table 5.2 indicated with a star are present. OPOS may therefore be viewed as a specialization of POS, focusing on specific behaviors for outdoor, multi-player games.

As OPOS is applied to each child separately and taking into account the number of categories of behaviors it is virtually impossible to perform coding in real-time, as is done with POS and SOPLAY. Instead, video recordings of children playing the targeted outdoor game are required for off-line analysis.

Table 5.3 Classes and subclasses of OPOS (continued on next page)

Class	Behavior	Explanation
Physical actvity	Intensive physical activity	Exhausting physical activity that one can not keep doing for a long period of time. For example: running, jumping or skipping.
	Non-intensive physical activity	Physical activity that one can keep doing for a longer period of time. For example: walking, moving arms or legs, bending and standing up, crawling, moving while staying on the same location, etc.
	No physical activity	Standing, laying or sitting still. Very small movements such as coughing yawning, putting your hands in your pocket, looking at your watch, etc. while being still should also fall in this category.
Focus	Looking at other players	The player is looking at one or more other players. This does not only include looking at the face, but also looking at other parts of the body.
	Looking at game objects	The player is looking at one or more game objects. All things that are part of the game besides players and surroundings are game objects. For example a ball, a goal, a chalked circle on the ground, a hand held object, a token, etc.
	Looking at something out of sight, possibly part of the game	Looking at objects, people or surroundings that are not part of the game.
	Looking at something else	When game objects or players are out of sight of the camera and the observed player is looking in the direction of which these player(s) or object(s) likely are.

Table 5.3 Classes and subclasses of OPOS (continued from previous page)

Class	Behavior	Explanation
Social interaction	Functional, with another player	All interactions (verbal en nonverbal) that are functional for playing the game and directed to one or more other players or to no-one. For example instructions such as "give me the ball!", "get the monster-coin!" and "tag him!", or expressions like "John is it!", "tag!" or counting points aloud, or physical contact such as tagging, holding hand, etc that are needed to play the game.
	Non-functional positive/ neutral, with another player	All interactions (verbal and nonverbal) that are not functional for playing the game, that are positive or neutral and directed to one or more other players or to no-one. For example communication about subjects that are not related to the game, showing results to other players, cheering, screaming, expressions of enjoyment and physical contact not required for playing the game such as holding hands, high five, etc.
	Non-functional negative, with another player	All interactions (verbal and nonverbal) that are not functional for playing the game, that are negative and directed to one or more other players or to no-one. For example negative communication such as swearing and bullying, expressions of pain or negative physical contact such as kicking or hitting.
	With a non-player	All interactions (verbal en nonverbal) that are directed to someone who is not a player in the game. This can be a researcher, a teacher, a parent, a peer who is watching the game, etc.
	Unintended physical contact	Physical contact that is not intended, such as accidentally bumping into another child.
General	In sight	In sight of the camera.
	Out of sight	Out of sight of the camera.

Procedure for applying OPOS

To apply OPOS to measure player behavior in outdoor game play, the following steps must be taken:

During the playtest

1. record the game play on video. Since children are likely to be moving around the play-area, we advise to use at least two cameras to capture as much of the gameplay as possible.

After the playtest

2. Review the videos. The behavior of *each* child is separately coded according to the classes of OPOS (see Table 5.3).

The classes physical activity and focus must be coded using a duration-based approach, i.e. the total duration (in seconds) of each behavior is calculated instead of counting the number of occurrences of the behavior. In contrast, the class 'social interaction' is coded using an event-based approach. Note that though it is possible to code duration-based behavior by hand, i.e. noting down all start and stop times, it is much faster to use dedicated software as this greatly speeds up both the coding process as well as the analysis afterwards; we used Noldus Observer* for this.

* http://www.noldus.com

3. Coding the observations results in a total duration (in seconds) of each subclass of physical activity, focus and general, for each child, e.g., child 1 showed 225 sec of intensive physical activity, 461 sec of non-intensive physical activity and 178 sec of no physical activity. For each subclass of social interaction it results in a number of occurrences of behavior of each child in one game, e.g., child 3 displayed 4 occurences of 'functional, with other player' behavior, etc.

4. Inevitably, games are seldom equal in duration. Therefore, to comparing the results of one game to another, for each child the durations are converted to a *percentage* of the total playtime. The event-recorded behaviors for social interaction are converted to an average per minute. The results of OPOS are these numbers.

5. The results for all children can now be further analyzed.

6. When multiple independent observers are involved in the analysis, it is possible to assess the reliability of the observations.

GROUPSORTER

Although it is important to gather quantitative evidence for evoked play behavior (which can be achieved through OPOS), observation as such does not offer insight into what players wish, feel, or how they evaluate their play experiences. For this purpose methods that elicit qualitative data from the participants are required to understand the game experience in general, but also about what elements of a game will evoke desired experiences (e.g., fun). This could give information to help improve existing games, as well as inform the design of new Head Up Games.

In the case of evaluating a multi-player game, we argue that it is most logical to question the children as a whole group since they have experienced the game together. This will likely encourage discussions between the players and therefore result in richer data. To interview a group at once, a common method to use is a focus group, or group interview (Krueger and Casey, 2000). Focus groups have been used extensively to gather children's views and opinions, e.g., Hennessy and Heary (2005). A group interview offers several advantages over an individual interview. It creates a more natural context for children to be interviewed: children feel more comfortable and relaxed in a peer setting. Furthermore, a greater openness and variety in responses can be expected (Eder and Fingerson, 2002), and there is less pressure on each child to respond to every question. Also, the focus group places the children in an expert position, which makes them feel less like being questioned and more as though they are sharing experiences with peers (Hennessy and Heary, 2005). Finally, interviewing children as a group may reduce the power imbalance that exists between the adult interviewer and the child interviewee (Eder and Fingerson, 2002). From a practical point of view doing group interviews is also convenient: interviewing all children together requires only one researcher to be present, and is easier to arrange to take place right after the game play, especially in school settings where our games are often evaluated. It takes up less time and planning for the teacher, compared to interviewing the children individually. Planning individual interviews consecutively can cause a long time lapse for the child between playing the game and being interviewed about it, which is less desirable.

Pitfalls that can occur in group interviews with children include intimidation or peer pressure that might inhibit participants to voice their own opinion; or to voice an opinion of another child, in a desire to fit in with the rest of the group (Hennessy and Heary, 2005). Even if such a desire is not present, a child might still repeat the opinion of another child, simply because he or she has listened to the other's opinion and readily agrees, before forming an opinion of his own (an effect known as 'cognitive tuning' (Fern, 2001)).

As has been pointed out by Hanna et al. (2004), ranking is a valid way for children to indicate their preference, as opposed to ratings. Ratings often have a "ceiling effect": children tend to give maximum ratings. An example of a ranking method specifically tailored for children is the Fun Sorter (Read, 2008). The Fun Sorter lets children rank items on one or more constructs, for example 'fun' or 'easy to use'.

We propose to combine rankings with the group interview to avoid some of the pitfalls mentioned above, as follows: first, we let the children rank elements of the game itself on the construct 'most fun'. The ranking is done individually. Subsequently, the ranking is repeated, but now as a group task in a focus group setting. However, the ranking of preferences is not per se the goal of the evaluation. While ranking in the group context, it is natural for children to discuss what element is ranked first, what element second, etc. and to provide arguments to their peers regarding their order of preferences. By recording the ensuing discussion, we hope to obtain qualitative insights into why children think items are fun, or not. By first having the children rank individually, we try to achieve that children properly form their own opinion first, and not merely follow the suggestions of a peer in the group discussion. Finally, since the items that the children are ranking automatically lead the discussion in that direction, no questioning route has to be prepared. Still, we expect that a moderator is needed to guide the discussion if children reach a dead end and cannot resolve an issue themselves, and to make sure that all children have an equal participation in the discussion.

Procedure for applying GroupSorter

Preparation

Save the Safe | p. 36

Choose items for ranking that are related to the game that is being evaluated. Note that the choice of these items will have a direct effect on the direction of the focus group discussion. For example, for *Save the Safe* the items included: "belt, running, ball, playing together".

After the playtest

1. Let the children rank the items individually. Note that it is advisable to not allow the children to sit side-by-side, to avoid copying of the rankings (see Figure 5.1).

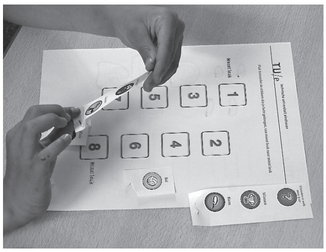


Figure 5.1 Step 1: A child ranking the items for Save the Safe from most fun to least fun

2. Repeat the ranking in a group setting. Record the ensuing discussion (see Figure 5.2).
3. Analyze results. We recommend not transcribing the data verbatim but only capture comments when they provide (1) an argument, more than just: "I liked it" or "I didn't like it", (2) an example of an experience or (3) a suggestion for improvement. Next, these comments can be analyzed using conventional content analysis (Hsieh and Shannon, 2005).

Figure 5.2 Step 2: A group of children ranking the items together

Concluding, the main result of this method will consist of qualitative comments that can be used to analyze the game play and experience.

CASE STUDIES

The methods described in the previous section have been applied during evaluation of three Head Up Games, which will be described in this section. For each case study we will shortly describe the game itself, if it has not appeared in this thesis before; otherwise we refer to the part of the thesis where the game first was described. Furthermore we will describe the larger context in which the study took place. Then we will focus on our experiences with OPOS and GroupSorter during the case study.

EVALUATING LIGHTHOUSE

Context and Setup

In contrast to the cases we describe next, where the focus was to evaluate a newly designed Head Up Game, this case study focused on the assessment of the reliability of OPOS. To that end, a game called *LightHouse* was developed in two versions: a game in which players competed individually against each other and one game in which players competed in teams.

One class of 24 children of a Dutch school participated in this experiment. Four groups of 5 to 7 children (age 10-11 years old) were formed; boys and girls were equally divided over the groups, though friends were grouped together to make the experiment more realistic. Two groups played the *LightHouse* game: one group played the individual version of the game, the other group the team version. The two other groups played tag and soccer respectively. The latter games were played and recorded to test whether OPOS was

LIGHTHOUSE

by Saskia Bakker

Figure 5.3 Children playing the LightHouse game.

LightHouse (Bakker et al., 2008) is a pirate-game in which players (the pirates) have to collect treasures (wooden coins) from treasure islands. The treasure islands are guarded by a lighthouse, emitting a rotating light from a desert island. If a player is 'touched' by the beam of the lighthouse, he loses all the coins he is carrying at that moment and has to return to his ship. Furthermore, at random times in the game, a seamonster attacks. To avoid the attack of the seamonster, the pirates have to get to the desert island as fast as possible. The last pirate to arrive, gets caught by the seamonster and half of his treasure is taken from him. One of the treasures on the island is the 'monster-coin'. The pirate who finds this treasure becomes friends with the seamonster, and as a consequence the seamonster will give this pirate the treasure that he captures from one of the other pirates. After this has happened the 'monster-coin' is returned to the islands, so other players can try and find it. The game ends when all treasures are gone from the islands. The pirate who captured the most treasures wins the game.

The islands are represented by chalked circles on the ground. A physical lighthouse object was built using a PIC processor. The lighthouse system is able to emit light and playback sounds and the playback of the sounds is operated by a remote control. It cannot detect the presence of a player in it's beam; this is achieved by a Wizard of Oz-setup: a researcher watches the game and operates the remote control if a player is touched by the beam. Subsequently, a sound is played from the lighthouse, to indicate to that player that he loses his coins. The arrival of the seamonster is indicated with a sound playing from the lighthouse, this too is operated by the researcher.

The LightHouse game can be played individually, or in teams of two. See Figure 5.3 for an impression of the LightHouse game.

capable of detecting differences in play behavior for games other than Head Up Games.

Each game was captured on video by one stationary camera. Then we applied OPOS to analyze the behaviors evoked by the traditional games (tag and soccer), and both versions of *LightHouse*. Also, we applied OPOS to a short video of two children playing a video game on an XBox, as totally different behaviors are displayed in such a game. Finally, we compared the results of OPOS in terms of the evoked physical activity and social interaction, as well as the visual focus of the players.

A secondary aim of the study was to compare the game experience of a technology enhanced (Head Up) game to a non-technological version of the same game. Therefore, the *LightHouse* game was redesigned in a non-technical adaptation of the original game. Two groups of children played both games and to compare the children's experiences GroupSorter was applied.

Experiences with OPOS

For the evaluation we have compared two versions of *LightHouse*, the individual and group version as described above, to three other games, namely tag, soccer, and a video game. The results of this comparison are shown in Figure 5.4, Figure 5.5 and Figure 5.6. Rather unsurprisingly, the video game scores very low for physical activity, in the other four outdoor games players are much more physically active. Furthermore, in the category 'focus' the game tag scores high for 'looking at other players' – this can be explained by the fact that there were no game objects to look at. Finally, comparing the other three outdoor games, the behavior 'looking at other players' occurred more in the *LightHouse* games than in the soccer game. Again, this can easily be explained: in the soccer game players focused most on the ball.

While all the data was coded by a single coder a selection of the data was coded by two additional coders as well (see also Bakker et al. (2008)) to calculate inter-coder reliability, resulting in the following Kappa coefficients: for physical activity K = 0,70 (acceptable agreement according to Landis and Koch (1977), for focus K = 0,45 (moderate agreement), and for social interaction K = 0,24 (low agreement). The

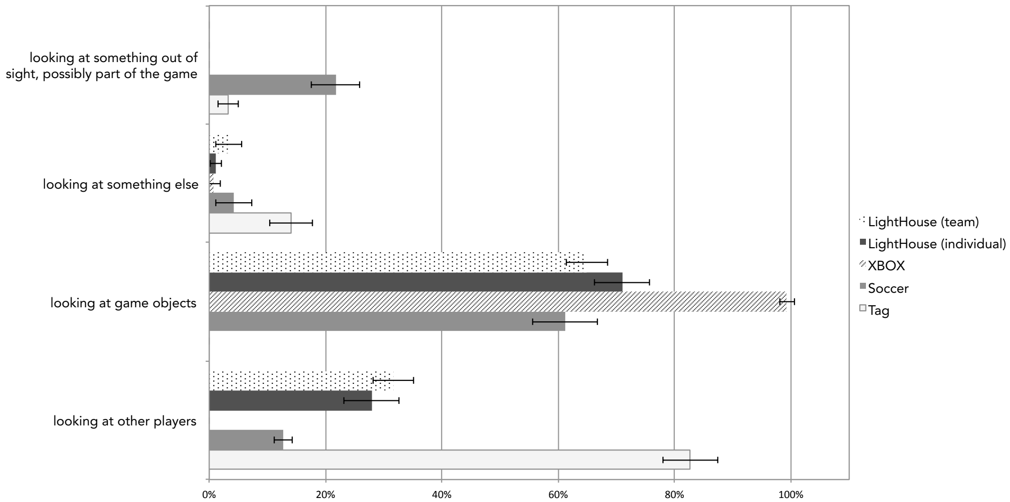


Figure 5.4 The results in the class 'focus'. Behaviors in average percentage of the total time.

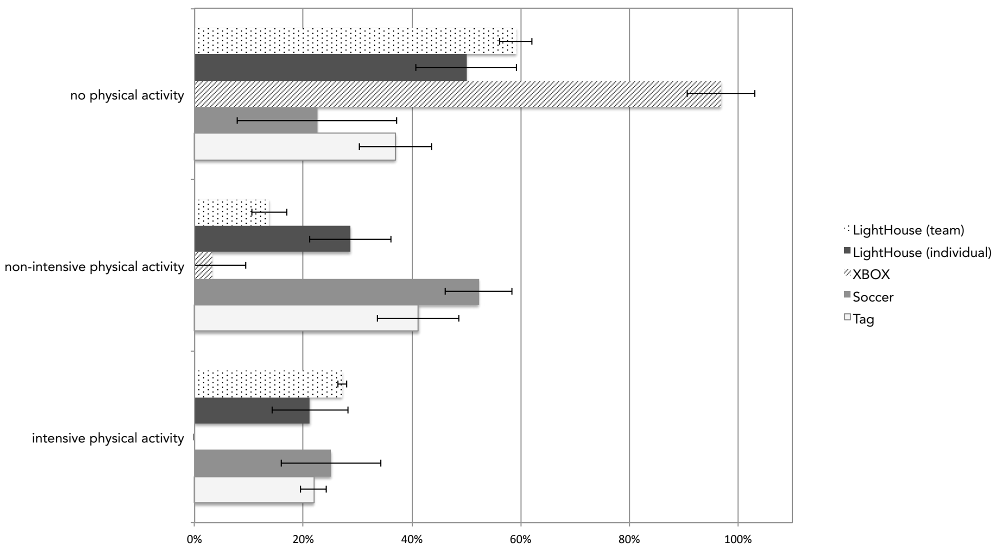


Figure 5.5 The results in the class 'physical activity'. Behaviors in average percentage of the total time.

low agreement in the class social interaction may have been caused by the way we recorded the games. The additional coders indicated that it was especially hard to interpret whether social interaction behaviors should be coded as functional or non-functional to the game as it was often difficult to understand exactly what the players were saying.

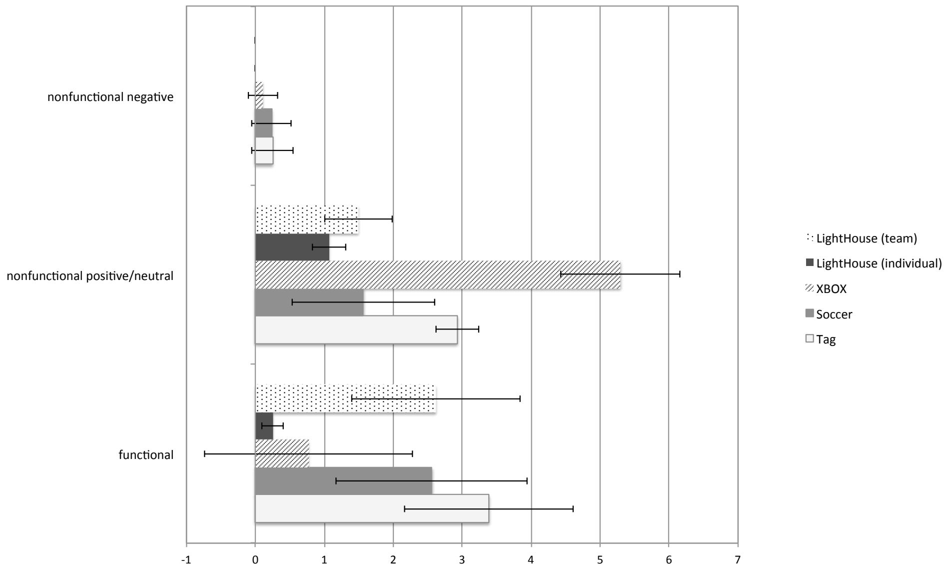


Figure 5.6 The results in the class 'social interaction'. Behaviors in average number of events per minute.

More sophisticated recording equipment such as multiple cameras could possibly overcome this problem.

Experiences with GroupSorter

Besides evaluating the applicability of OPOS, we examined the potential added value of technology and its potential downsides by comparing the technology-enhanced version of the *LightHouse* game to the non-technological version of the *LightHouse* game, i.e. without technological enhancements. In the technology-enhanced version of *LightHouse*, players had to collect treasures without being seen by a 'physical' lighthouse. This game object emits a rotating light and makes a horn-sound when a player is 'seen'. In the non-technological version, the lighthouse is acted out by one of the researchers, who created the rotating light with a torch and screamed in case a player is 'seen'. All other artifacts and rules are the same for both games.

Two groups of 6 children (age 10-11) played both the technology-enhanced and the non-technological version of the *LightHouse* game, and did a GroupSorting exercise afterwards to enable qualitative analysis of the game experience of both games. For the GroupSorter, we visualized four major activities of the game on cards. Each activity was represented on two cards, one showing the technological version of *LightHouse*, and one showing the non-technological version. We first used these cards to explain the GroupSorting procedure, after which we gave each child a sheet of paper showing all eight cards (the order of the cards was systematically varied to prevent order effects). The children were asked to make their individual ranking on these sheets, and then come together to make the group rating using the large cards. A researcher moderated the discussion during the group session, by frequently asking why certain decisions had been made.

As intended, the discussions focused mainly on the differences between the technology-enhanced and the non-technological version of the game. Having the large cards at hand enabled the children to clearly distinguish and discuss both versions of the game. The discussions were recorded on video and later analyzed for quotes revealing differences between the two versions. In total, close to 30 quotes were captured which were analyzed through conventional content analysis (Hsieh and Shannon, 2005). The general opinion was that the technology-enhanced version was more exciting or 'cool'.

The moderators in this study felt that the discussions were sometimes influenced by peer pressure: one child would make a suggestion and the others would agree without discussion. In such cases, the moderators tried to stimulate discussion by asking specific participants for their thoughts. The fact that children already had their personal rating at hand, enabled the moderator to ask specific questions (e.g., 'you had a different preference in your personal rating, why is that?') to encourage discussion. This may have been harder when 'traditional' focus group techniques would have been used. One group-discussion was moderated by the researcher who supervised the technology-enhanced version of *LightHouse* and the other by the supervisor of the non-technological version. We expected that children might argue in favor of the game that their interviewer had supervised. However, when comparing the opinions of both groups, this does not seem to have influenced the results.

EVALUATING SAVE THE SAFE

Context and Setup

The aim of this study was to investigate the effect of a tangible versus a virtual game element on the game experience. To this end, we designed two versions of the *Save the Safe* game. 27 children (8-9 years) played both versions of the game, in four groups of eight children per session. Genders were evenly distributed over the sessions, and the order of play was systematically varied over the sessions. Each session started with a training game, then both versions of the game were played, and for the final game the children decided which version of the game they wanted to play – similar to the AgainAgain method (Read and Macfarlane, 2006), where children are asked in a survey what product they would want to use again. All sessions were videotaped using two cameras at corners of the play field. Also, two children were equipped with head mounted cameras. Each group participated in a GroupSorter session after playing both versions of the game. The evaluation was conducted at a school, where children from one class participated. The evaluation took place over the course of one school day.

Save the Safe | p. 36

Experiences with OPOS

Based on the experiences during the evaluation of the *LightHouse* game, we decided to apply OPOS unchanged in the evaluation of *Save the Safe*. Despite the low inter-coder reliability in the social interaction class found during the *LightHouse* evaluation case, we believe it is valuable to distinguish functional and non-functional social interaction, in order to allow more detailed comparison between different games. To capture social interactions more clearly, we decided to videotape the sessions by two stationary cameras placed on two corners of the playing field, which was a school playing field.

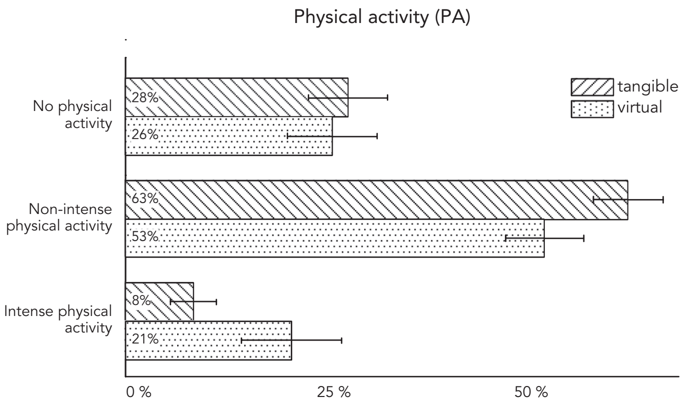


Figure 5.7 The results in the class 'physical interaction'. Behaviors in average percentage of total time, with 95% confidence intervals.

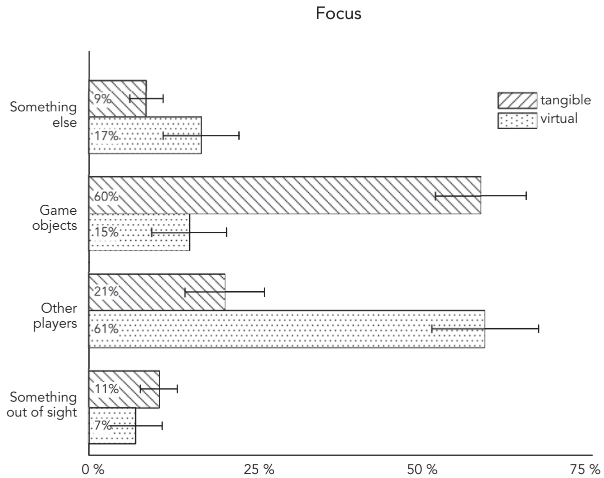


Figure 5.8 The results in the class 'focus'. Behaviors in average percentage of total time, with 95% confidence intervals.

	Type		Order		Interaction	
	F(1,22)	p	F(1,22)	p	F(1,22)	p
Functional	0.536	0,472	0,249	0,623	0,021	0,886
Non-player	4.789	0.040*	1.535	0.228	2.702	0.114
Positive [a]						
Negative	1.565	0.224	0.634	0.434	8.062	0.010

*significant at 5% level.

[a] too few observations.

Table 5.4 The results in the class 'social interaction'.

The results of the comparison between the tangible and virtual version of the game are shown in Figure 5.7, Figure 5.8 and Table 5.4. There is a significant difference between the percentage of intensive physical activity between the two games. Thus, in the game with the virtual element, children engaged in more intensive physical activity than in the game with the tangible element. Furthermore, from the results in the category 'focus', we can conclude that the children looked much more at other players in the virtual game than in the tangible game. Vice versa, children looked more at game objects in the tangible game, which can be easily explained since a ball was used in that game.

All data was coded by a first observer, while 25% of the data was coded by a second observer to calculate inter-rater reliability. This resulted in Kappa coefficient K = 0,53 for the category physical activity, and K = 0,43 for the category focus. According to Landis and Koch (1977) these values indicate moderate agreement between coders. For social interaction K = -0,24, indicating no agreement. We must conclude that using multiple cameras did not make it easier to code social interaction behaviors. The footage of the head mounted cameras was not used in this analysis. We had some trouble with the fitting of the helmets; the turned out to be too loosely fitted to the children's heads, resulting in unusable footage.

Experiences with GroupSorter

We applied the GroupSorter method in order to find out how different elements of the game contributed to the overall game experience.

The items that we asked the children to rate from most fun to least fun were: the belt, the safe, running, discussing, playing together, the ball, the team, and winning. Note that the elements of both games are mixed together to create eight items for one ranking. This allows us to compare the elements in relation to each other; if we would have the children rate the elements in a ranking per game, we would not be able to do this. We presented the items as text with icons on stickers that children could stick on a piece of paper.

Children seemed to have no problems understanding the concepts they were asked to rank, and they could perform the task quite easily. Immediately after the individual ranking the group discussion took place. The discussion was recorded and we analyzed it using conventional content analysis (Hsieh and Shannon, 2005).

All items that were on the list to be ranked were mentioned in the discussion, although some items were discussed more intensively than others. This mainly happened when participants did not agree on an item. The ensuing discussion gave much insightful information about the reasons for liking or disliking a particular element of the game. No new items were mentioned in the discussion, but the discussion did reveal some unexpected game dynamics of game elements influencing each other in a way that we had not foreseen when designing the game. Finally, the things that were said during the discussion were in agreement with the rankings the children had given. For example, many positive remarks were given on the belt, which scored high in the ranking.

As opposed to the evaluation of *LightHouse*, in the evaluation of *Save the Safe* moderators did not note occurrences of peer pressure; on the contrary, children felt apparently quite at ease voicing a different opinion than their peers, and were able to work together to make a ranking, as the following discussion shows:

for the sake of clarity, the names of the items that the children were ranking are underlined

child 1: I thought <u>playing together</u> was most fun

child 2: [on my ranking] I have the <u>belt</u> on one, and <u>winning</u> on two

child 3: oh, but I find <u>winning</u> is not so important

child 2: shall we put the <u>belt</u> here then [points to a spot on the ranking chart]?

[another child starts to put items on the chart without consulting the others]

child 4: Robert, don't fill it in yourself, we have to agree first!

Although it might be difficult to extract from the above transcript, from the audio it was clear that children were able to freely voice their opinions and discussed to come up with a common agreement to the ranking. Though one child made an attempt to impose his own ranking on the rest, he was immediately corrected by his peers, no intervention from the moderator was necessary. However, we noticed that a moderator is absolutely necessary. In one particular case a child started bullying another child, as they had contrasting opinions. In that case the moderator intervened to correct the behavior of the children. Furthermore, in contrast with 'normal' focus groups, the moderator did not have a set of guiding questions. The items that had to be ranked automatically provided guiding in the discussion. Most of the time when children did not agree on where to rank an item, a discussion would inevitably follow. However, on items that the children were in perfect agreement, comments such as "yes, that's it" were often given without further explanation. The moderator would then prompt the children to elaborate on their choice. The transcript below shows this:

mod: Which game did you like best?

child 1: I liked the game with the belt most

[other children agree]

mod: Does everybody agree? [child2] What do you think?

child 2: No, I found the game with the ball the most fun

[… more discussion on which game was most fun...]

mod: Why did you find the game with the belt the most fun?

child 3: I liked the sensation of the belt, and the fact that I had to run faster [compared to the game with the ball]

EVALUATING HEARTBEAT

Context and Setup

In this case study we incorporated biofeedback in an outdoor game. The game was developed in an iterative process, involving children early on in the process. The resulting game was playtested with 32 children, ages 11 to 13. The aim of the evaluation was to compare two versions of the game, one incorporating heart rate measurement and one not. Participants played the game in groups of 8 players and each player experienced both versions. The game play was recorded and the video footage was coded using OPOS. Two classes played the game during a school day. The game was played in a park near their school and evaluations were conducted in the school. After playing both versions of the game, GroupSorter was applied to gather the children's feedback.

Experiences with OPOS

The game of *HeartBeat* was played in a large outdoor park, with bushes and trees for the children to hide behind, making it impossible to capture the play-behavior with one or two stationary cameras, as has been done in the game evaluations of *LightHouse* and *Save the Safe*. To solve this issue, two players in each game received a head-mounted camera. This made it possible to review the play behavior from one player's point of view.

HeartBeat | p. 37

Save the Safe | p. 36
LightHouse | p. 157

From the play sessions 10 videos were selected for observation, 5 videos of the game played with heart rate monitor and 5 videos of the game played without the heart rate monitor. One observer coded the videos, using only the categories of physical activity and social interaction. The category 'general' was irrelevant, since the camera was head-mounted. The category of focus was left out, since it was impossible to determine exactly what the player was looking at from the head-mounted camera footage (e.g., a player looking down can be watching his footsteps to see where he is running, but can alternatively be watching his game device, or can try to stay hidden for other players). A selection of the material was also coded by a second coder, and from this the inter-rater reliability was calculated: for the category physical activity this resulted in the Kappa coefficient K = 0,90, and for the category social interaction Kappa coefficient was calculated at K = 0.73. Our main concern in coding the footage was to prove that our implementation of heart rate monitoring did not have a negative effect on physical activity. Since the difference between the two game versions was minimal, the results from OPOS did not show significant differences.

Experiences with GroupSorter

In the evaluation of *HeartBeat* the same approach as in *Save the Safe* was used. Children sorted five game elements according to how much 'fun' they experienced them. The game elements referred to various physical, social and technical game elements: tagging other players, random team allocation or 'hearing' other players through your game device. Four focus groups were held, three of them with 8 participants, one with 3 participants. The average length of a focus group session was 15 minutes. Directly after playing two versions of the game, the children were taken into an empty classroom where the discussions were held. Children first filled out their own ranking scheme and subsequently discussed it with others.

In our experience the atmosphere in the focus groups was mixed. Two moderators each conducted two sessions with different groups. There was a noticeable difference in atmosphere between the two moderators. One moderator had very calm sessions, where children waited for each other and generally agreed upon most topics. The other moderator had sessions with more discussions. Children disagreed with each other and felt free to speak aloud to describe their own experience or opinion. When reviewing the focus group footage we experienced a fairly balanced distribution of speaking turns. Although more assertive children obviously accounted for more comments, the individual ranking scheme allowed the moderators to mingle silent children into the discussion. Only in very few cases these silent children agreed to everything. We also observed that the discussion exclusively revolved around the chosen game elements. Children did not bring other topics into the discussion.

In general the GroupSorter exercise provided us with numerous examples of game situations. Children mostly argued why they liked, or didn't like game elements based on what they experienced. For example, one child reported she didn't like a technical feature, because it failed several times during play. When the moderator inquired whether she would have liked the feature when it would not have failed, she agreed to this.

The moderators were given a set of guidelines and back-up questions. One very important rule in the discussion was that only the moderator was allowed to place or change the order of the cards. Children first had to convince others that a game element should be ranked higher or lower, before the moderator would change it.

REFLECTIONS

OPOS

OPOS provides a concrete and operational way to describe salient aspects of outdoor play, more specifically for Head Up Games. This is a new class of games and as far as we are aware there are no existing observation schemes that measure social interaction and physical activity conjointly for outdoor, multiplayer games.

Figure 5.9 One of the children wearing a head-mounted camera during Save the Safe

Obtaining data, in our case video footage, is not as straightforward as in laboratory-based evaluations. Players tend to disperse and move fast. Depending on the game and the play area multiple cameras could be used; alternatively individual players may be tracked and coding play behaviors may then be done for individual players. In the evaluation of *Save the Safe* and *HeartBeat*, some players were equipped with wireless motion cameras that were mounted on helmets (see Figure 5.9). This did allow us to capture the play behavior of one player for the whole duration of the game, in contrast to stationary cameras where children would sometimes disappear out of sight. Also, because the head mounted camera was much closer to the player it did a better job at capturing the 'feel' of the play. On the other hand, the footage it provided gave a limited view of the playing field as it captures the perspective of a single player only, compared to a stationary camera, which can capture multiple players simultaneously. Sometimes the footage of the head-mounted cameras was ambiguous, as it was not always clear what the player was focusing at (see for example Figure 5.10). Summarizing, there is no solution that is best for every situation; we advice to select the type of camera (head-mounted or stationary) taking into account the outdoor environment in which the game is to be evaluated.

Figure 5.10 Screenshot of footage of a head-mounted camera. What is this player looking at: another player or the ball?

It is worth mentioning that applying OPOS is a very time-consuming activity. The footage has to be reviewed for each player individually and for each player three classes of behavior need to be coded. For example, a game of eight players of five minutes would take approximately 8x3x5 = 120 minutes of coding.

Our experiences with applying OPOS have shown that it is relatively easy to code the classes 'physical activity' and 'focus'. However, in our studies of *LightHouse* and *Save the Safe*, inter-rater reliability was low for the class 'social interaction'. The main reason for this is that it was very hard to judge from the footage of stationary cameras whether the children were engaged in social interaction, since the children's faces were not properly discernible. Also, their verbal utterances were not very audible, because the children would often be too far away from the camera, making it difficult to distinguish positive and negative interactions. As a direct consequence, the results for 'social interaction' are not reliable. In contrast, in the study of *HeartBeat* the inter-rater reliability of coding the class social interaction was higher. This is due to the fact that the footage was captured by the head mounted cameras and therefore much better recorded what the children were saying. We think that another way to improve capturing

Save the Safe | p. 36
LightHouse | p. 157

HeartBeat | p. 37

social interaction is to separately capture audio recordings for analysis. To this end, every player could be equipped with small audio recording devices. This would also solve the problem of children running in and out of sight of the camera, at least for the class 'social interaction'.

In related work on measuring physical activity another means of data capturing that is often referred to is deploying accelerometers (e.g., Leal Penados et al. (2009)), though these can be expensive (Dollman et al., 2009). Using an accelerometer would probably result in even more accurate measurements of the levels of physical activity, but this would come at the cost of losing the rich contextual information that is provided when using observational data (Loprinzi and Cardinal, 2011). Though data from accelerometry would speed up the process of analyzing physical activity levels, we would still argue to combine it with observational data, to correlate the levels of physical activity with interesting events in the game play.

GROUPSORTER

Overall, we conclude that GroupSorter can provide valuable feedback on the game during the design process of a game. GroupSorter is specifically designed for interviewing groups of children, while avoiding common problems that are known to arise when doing so. It provides both quantitative as well as qualitative feedback: the quantitative result consists of a ranking of elements of a game on the construct 'fun'. The most valuable data from GroupSorter, though, is the qualitative data: the method elicits comments of children regarding the game play and experience that designers can use to improve their product in the next design cycle of the development. However, there are a number of issues in general that need attention when applying this technique, which we will discuss here.

First, the items to be ranked must be picked carefully, because they will to a great extent influence the direction of the group discussion. From the analysis we conclude that the discussions will remain centered around the items provided to the children; new items did not surface in the discussions. Therefore, aspects of the game that are of interest for the evaluation should be picked as items for the ranking, to ensure that these aspects are discussed.

In our evaluations we used icons as well as text to represent each item. This clarified the items to be ranked; from the focus group discussions we did not get the impression that children had misunderstood or misinterpreted the items. We would like to underline the importance of clear visualizations of items to be ranked, especially when multiple items are alike and may be confused by the participants (such as in the *LightHouse* study, where we compared a technology-enhanced and a non-technical version of the same game). Large cards may support the group discussion in such cases.

Though the use of items to guide the discussion will lessen the tendency of children to give “desired” answers (Eder and Fingerson, 2002; Garbarino and Stott, 1992), and no interview questions need to be prepared, still the moderator needs to take care that he prompts the children for explanations of their rankings. Often, while the group is ranking, explanations are offered automatically, but sometimes children would immediately agree on an item, and no further details would be given. At such a moment, the moderator prompted the children to elaborate on their ranking, to obtain more insight.

We had no trouble annotating the children’s discussion from the audio recording, although at some times children will talk at the same time. A way to deal with this is to record each child’s input individually, by equipping each child with a personal microphone, as for example Sluis-Thiescheffer et al. (2007) have done.

A possible pitfall of a group based survey technique is that some children might be dominated by others, and their opinions might be unvoiced or dismissed quickly (Lewis, 1992). We also experienced this in the *LightHouse* study, where we felt that peer pressure had influenced the focus group session. Clearly, more vocal and assertive children can influence the final group sort more but it is up to the moderator to help all children voice their thoughts and feel that their individual opinion is valued. For example, the moderator can address a child directly to invite him into the conversation, while silencing the more vocal kids. Also, as we expected, individually ranking the items before the group discussion opens the opportunity to invite children into the discussion. By comparing their ranking to the group-ranking scheme, the moderator can ask why a child answered differently. Finally, because children have already made their ranking individually, there is less chance of the effect called "cognitive tuning" (Fern, 2001), where children could be influenced by other children's remarks, before making up their own mind about the items. Particularly in the *Save the Safe* and *HeartBeat* case study, we have experienced that referring to the children's individual rankings during the focus group encourages children to formulate their own opinions.

DISCUSSION & CONCLUSION

In this chapter we have described two novel methods, developed for the evaluation of Head Up Games. We have described OPOS, an observation scheme that quantifies outdoor play behaviors in which we are interested within the scope of Head Up Games. GroupSorter is a method for gathering subjective data from groups of children. In the previous section we have reflected on the use of both methods in three case studies. Here we discuss the applicability of the methods in other contexts and participants.

APPLYING THE METHODS IN OTHER CONTEXTS

The methods described have been used to evaluate outdoor games. With respect to the applicability of these methods outside game development we can say the following: as the observation scheme OPOS is a strongly linked with outdoor play, its use beyond play or outdoor related applications is probably limited. In contrast, the GroupSorter method is a much more generic method and can be used for any kind of evaluation that involves multiple children. Also, as we have shown it generates both quantitative and qualitative data, and – in case the data is in concordance, the quantitative data can substantiate the qualitative data, and vice versa, providing a more solid proof for the found results.

APPLYING THE METHODS FOR OTHER AGE GROUPS

Both methods have been used for children in the age range of 7 to 13 years old, and have been found appropriate for this age group. However, since OPOS measures behavior only, without requiring direct input from the users, we argue that it is equally usable for other ages, both younger and older. Finally, the GroupSorter method is similar to a focus group, and we deduct from literature that focus groups can be successfully applied for ages six and up (Hennessy and Heary, 2005).

HOW TO INTERPRET THE RESULTS OF OPOS?

By using OPOS, objective quantitative data is gathered about the play behavior that emerges during game play. However, from a design perspective, one should keep in mind that in this case "more" is not always "better". For example, in our vision of Head Up Games we propagate the aspect of physical activity in outdoor play. Creating a game that will enforce players to physically exhaust themselves (and thus generating a high value on the 'physical activity' scale of OPOS) does not automatically result in a game that is most fun.

WHEN TO USE GROUPSORTER OR OPOS

The two methods discussed in this chapter provide insights on different levels. Regarding the use of each method we make the following remarks:

GroupSorter elicits qualitative feedback from children on the game and the game experience. As such the method can very well be used formatively in a game design process, in which children are involved as informants for designers.

In contrast, OPOS results in quantitative data of the game behaviors, and we argue that it is best used in a summative assessment, for example OPOS can help quantify differences between play behavior of different (versions of) games.

In its current form, OPOS is too time consuming to apply during a design process. Reviewing the footage for each child simply takes too long. To reduce the time investment, an interesting line of research would be to see whether or not (some of) the data can be captured automatically, for example using the game technology. For example, the RaPIDO game devices could record movement, and could capture children's speech for analysis of the social interaction (though the latter would still require time to analyze after the evaluation).

REFERENCES

Anderson, C. A., and Carnagey, N. L. (2009). Causal effects of violent sports video games on aggression: Is it competitiveness or violent content? Journal of Experimental Social Psychology, 45(4), 731–739.

Bakker, S., Markopoulos, P., and Kort, Y. de. (2008). OPOS: an observation scheme for evaluating head-up play. In Proc. of NordiCHI '08 (pp. 33–42). Presented at the NordiCHI'08, ACM, Lund, Sweden.

Barendregt, W., Bekker, M. M., and Baauw, E. (2007). Development and evaluation of the problem identification picture cards method. Cognition, Technology & Work, 10(2), 95–105.

Bekker, T. M., and Eggen, B. H. (2008). Designing for children's physical play. In CHI '08 Extended Abstracts on Human Factors in Computing Systems (pp. 2871–2876). ACM, Florence, Italy.

Bekker, T., and Sturm, J. (2009). Stimulating physical and social activity through open-ended play. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 309–312). ACM, New York, NY, USA.

Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Rowland, D., and Benford, S. (2006). Interweaving mobile games with everyday life. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 417–426). ACM, Montréal, Québec, Canada.

Benford, S., Crabtree, A., Flintham, M., Drozd, A., Anastasi, R., Paxton, M., Tandavanitj, N., Adams, M., and Row-Farr, J. (2006). Can you see me now? ACM Trans. Comput.-Hum. Interact., 13(1), 100–133.

Benford, S., Rowland, D., Flintham, M., Drozd, A., Hull, R., Reid, J., Morrison, J., and Facer, K. (2005). Life on the edge: supporting collaboration in location-based experiences. In Proc. of CHI '05 (pp. 721–730). ACM, Portland, Oregon, USA.

Boren, M. T. (1999). Conducting Verbal Protocols in Usability Testing: Theory and Practice. Univ. Washington, Seattle.

Caillois, R., and Barash, M. (2001). Man, play, and games. University of Illinois Press.

Chen, V., Duh, H., Phuah, P., and Lam, D. (2006). Enjoyment or Engagement? Role of Social Interaction in Playing Massively Mulitplayer Online Role-Playing Games (MMORPGS). In R. Harper, M. Rauterberg, and M. Combetto (Eds.), Entertainment Computing - ICEC 2006 (Vol. 4161, pp. 262–267).

De Kort, Y. A. W., IJsselsteijn, W. A., and Poels, K. (2007). Digital games as social presence technology: Development of the social presence in gaming questionnaire (SPGQ). In Proc. of PRESENCE 2007: The 10th Int. Workshop on Presence (pp. 195–203).

Dollman, J., Okely, A. D., Hardy, L., Timperio, A., Salmon, J., and Hills, A. P. (2009). A hitchhiker's guide to assessing young people's physical activity: Deciding what method to use. Journal of Science and Medicine in Sport, 12(5), 518–525.

Donker, A., and Markopoulos, P. (2002). A comparison of think-aloud, questionnaires and interviews for testing usability with children. In PEOPLE AND COMPUTERS (pp. 305–316). Presented at the HCI.

Eder, D., and Fingerson, L. (2002). Interviewing Children and Adolescents. In J. F. Gubrium and J. A. Holstein (Eds.), Handbook of Interview Research : Context & Method. SAGE.

Fern, E. F. (2001). Advanced focus group research. Sage Publications Inc.

Fullerton, T., Swain, C., and Hoffman, S. (2004). Game design workshop. Focal Press.

Garbarino, J., and Stott, F. M. (1992). What Children Can Tell Us: Eliciting, Interpreting, and Evaluating Critical Information from Children. Jossey-Bass.

Hanna, L., Neapolitan, D., and Risden, K. (2004). Evaluating computer game concepts with children. In Proc. of the 2004 Conf. on Interaction Design and Children: Building a Community (pp. 49–56). ACM, Maryland.

Hassenzahl, M. (2011). User Experience and Experience Design. (M. Soegaard and R. F. Dam, Eds.)Interaction-Design.org, 1–14.

Haug, E., Torsheim, T., Sallis, J. F., and Samdal, O. (2010). The Characteristics of the Outdoor School Environment Associated with Physical Activity. Health Education Research, 25(2), 248–256.

Hennessy, E., and Heary, C. (2005). Exploring children's views through focus groups. In Researching Children's Experience: Approaches and Methods (pp. 236–252). Sage Publications Ltd.

Hetherington, E. M., Cox, M., and Cox, R. (1979). Play and Social Interaction in Children Following Divorce. Journal of Social Issues, 35(4), 26–49.

Hsieh, H.-F., and Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. Qual Health Res, 15(9), 1277–1288.

Huizinga, J. (1955). Homo ludens: A study of the play-element in culture. Beacon Press.

Jegers, K. (2008). Investigating the Applicability of Usability and Playability Heuristics for Evaluation of Pervasive Games. In Internet and Web Applications and Services, 2008. ICIW '08. Third Int. Conf. On (pp. 656–661).

Kirriemuir, J., and McFarlane, A. (2004). Literature review in games and learning. Futurelab. Retrieved from http://telearn.archives-ouvertes.fr/docs/00/19/04/53/PDF/kirriemuir-j-2004-r8.pdf

Krueger, R. A., and Casey, M. A. (2000). Focus groups a practical guide for applied research. SAGE.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–74.

Lantz, F. (2007). *PacManhattan*. In F. Borries, S. P. Walz, and M. Böttger (Eds.), Space Time Play (pp. 262–263). Birkhäuser Basel. Retrieved from http://dx.doi.org/10.1007/978-3-7643-8415-9_94

Leal Penados, A., Gielen, M., Stappers, P.-J., and Jongert, T. (2009). Get up and move: an interactive cuddly toy that stimulates physical activity. Personal and Ubiquitous Computing, 14(5), 397–406.

Lewis, A. (1992). Group Child Interviews as a Research Tool. British Educational Research Journal, 18(4), 413.

Lindley, S. E., Le Couteur, J., and Berthouze, N. L. (2008). Stirring up experience through movement in game play: effects on engagement and social behaviour. In Proceeding of the Twenty-sixth Annual SIGCHI Conf. on Human Factors in Computing Systems (pp. 511–514). ACM, New York, NY, USA.

Loprinzi, P. D., and Cardinal, B. J. (2011). Measuring children's physical activity and sedentary behaviors. Journal of Exercise Science and Fitness, 9(1), 15–23.

Markopoulos, P., Read, J. C., Macfarlane, S., and Höysniemi, J. (2008). Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers. Morgan Kaufmann.

McDonald, S., Edwards, H. M., and Zhao, T. (2012). Exploring Think-Alouds in Usability Testing: An Int. Survey. IEEE Transactions on Professional Communication, 55(1), 2–19.

McKenzie, T. L., Marshall, S. J., Sallis, J. F., and Conway, T. L. (2000). Leisure-Time Physical Activity in School Environments: An Observational Study Using SOPLAY. Preventive Medicine, 30(1), 70–77.

Mellou, E. (1994). Play Theories: A contemporary review. Early Child Development and Care, 102(1), 91–100.

Metin, P. (2003). The effects of traditional playground equipment design in children's developmental needs. MIDDLE EAST TECHNICAL UNIVERSITY.

Mueller, F., Agamanolis, S., and Picard, R. (2003). Exertion interfaces: sports over a distance for social bonding and fun. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 561–568). ACM, Ft. Lauderdale, Florida, USA.

Norman, D. A., and Draper, S. W. (1986). User Centered System Design: New Perspectives on Human-computer Interaction (1st ed.). CRC Press.

Nova, N., Girardin, F., Molinari, G., and Dillenbourg, P. (2006). The Underwhelming Effects of Automatic Location-Awareness on Collaboration in a Pervasive Game. In Int. Confernce on Cooperative Systems Design (COOP 2006) (pp. 224–238). Retrieved from http://infoscience.epfl.ch/record/85665

Parkes, A. J., Raffle, H. S., and Ishii, H. (2008). Topobo in the wild: longitudinal evaluations of educators appropriating a tangible interface. In Proc. of the Twenty-sixth Annual SIGCHI Conf. on Human Factors in Computing Systems (pp. 1129–1138). ACM, New York, NY, USA.

Parten, M. B. (1932). Social participation among pre-school children. The Journal of Abnormal and Social Psychology, 27(3), 243–269.

Poels, K., IJsselsteijn, W. A., and Kort, Y. A. . (2008). Development of the Kids Game Experience Questionnaire: A self report instrument to assess digital game experiences in children. In Poster Presentation at the Meaningful Play Conf. in Michigan.

Read, J. C. (2008). Validating the Fun Toolkit: an instrument for measuring children’s opinions of technology. Cogn. Technol. Work, 10(2), 119–128.

Read, J. C., and Macfarlane, S. (2006). Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In Proc. of IDC ’06 (pp. 81–88). ACM, Tampere, Finland.

Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., Harris, E., Smith, H., Randell, C., Muller, H., O’Malley, C., Stanton, D., Thompson, M., and Weal, M. (2004). Ambient wood: designing new forms of digital augmentation for learning outdoors. In Proc. of IDC ’04 (pp. 3–10). ACM, Maryland.

Rogers, Yvonne, Scaife, M., Harris, E., Phelps, T., Price, S., Smith, H., Muller, H., Randell, C., Moss, A., Taylor, I., Stanton, D., O’Malley, C., Corke, G., and Gabrielli, S. (2002). Things aren’t what they seem to be: innovation through technology inspiration. In Proc. of the 4th Conf. on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (pp. 373–378). ACM, London, England.

Rubin, K. H. (2001). The Play Observation Scale (POS). Center for Children, Relationships and Culture, University of Maryland. Retrieved from http://www.rubin-lab.umd.edu/Coding%20Schemes/

Rubin, K. H., Watson, K. S., and Jambor, T. W. (1978). Free-Play Behaviors in Preschool and Kindergarten Children. Child Development, 49(2), 534.

Salen, K., and Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals (illustrated edition.). The MIT Press.

Scarlett, W. G., Naudeau, S. C., Salonius-Pasternak, D., and Ponte, I. C. (2004). Children's Play (1st ed.). Sage Publications, Inc.

Sherry, J. L. (2001). The effects of violent video games on aggression. Human Communication Research, 27(3), 409–431.

Sluis-Thiescheffer, W., Bekker, T., and Eggen, B. (2007). Comparing early design methods for children. In Proc. of the 6th Int. Conf. on Interaction Design and Children (pp. 17–24). ACM, Aalborg, Denmark.

Soler-Adillon, J., and Parés, N. (2009). Interactive slide: an interactive playground to promote physical activity and socialization of children. In Proc. of the 27th Int. Conf. Extended Abstracts on Human Factors in Computing Systems (pp. 2407–2416). ACM, Boston, MA, USA.

Verhaegh, J., Fontijn, W. F. J., Aarts, E. H. L., and Resing, W. C. M. (2012). In-game assessment and training of nonverbal cognitive skills using TagTiles. Personal and Ubiquitous Computing, Special Issue on Child Computer Interaction, in press.

Vermeeren, A. P. O. S., Den Bouwmeester, K., Aasman, J., and De Ridder, H. (2002). DEVAN: A tool for detailed video analysis of user test data. Behaviour & Information Technology, 21(6), 403–423.

Chapter 6

Conclusion

SUMMARY OF THE WORK

Camelot | p. 34

In 2005 the design and creation of *Camelot* inspired a genre of games of its own; up until then pervasive, outdoor games for children were mostly location-based, smartphone-centered affairs. Though the aims of these games might have included enhancing the social interaction of children, in reality, because of the technology deployed, the children played these games most of the time *head down*, independently, engrossed by the visual interaction taking place on their portable screens. In contrast, *Camelot* aimed to make more prominent play behaviors seen in traditional outdoor games, like tag and hide-and-seek. Rich social interaction and physical activity are the main components of these outdoor games, and with *Camelot* we made a first attempt in creating a game that truly supported these behaviors, instead of interfering with them.

We described the vision of Head Up Games in Chapter 2 and illustrated it with several design cases. Once the vision of Head Up Games stood firm, we started to explore how to operationalize or even measure the "head upness" of a game. For this we decided to focus on the intended play behavior, as seen in traditional games. Bakker et al. (2008) created for this purpose the Outdoor Play Observation Scheme. Furthermore, we acknowledged that not only the play behavior was indicative for a game; we found the play experience of a game equally important. GroupSorter was created to gather qualitative player feedback on the play experience. GroupSorter provided a framework to interview a group of children simultaneously. There was a need for such a method too: in practice we evaluated our games with as many as 30 children in a few hours. In such settings it is virtually impossible to interview each child separately. OPOS and GroupSorter were applied for evaluating three Head Up Games, which we reported on in Chapter 5. Regarding OPOS, it must be noted that it requires a considerable time-investment; in Chapter 5 we identified possible solutions for reducing the time needed. GroupSorter proved to be an effective way to interview a group of children at the same time.

Save the Safe | p. 36
HeartBeat | p. 37
LightHouse | p. 157

Save the Safe, *HeartBeat*, and *LightHouse* were created after *Camelot* according to the vision of Head Up Games. For every game, the technology was implemented from the ground up, a time consuming and, sometimes, costly process and one that diverts the focus of attention from interaction and game design to engineering aspects. Creating these games gave us valuable insight into what technologies and interaction styles were appropriate for Head Up Games and based on our experiences we decided to build a platform that included this technology, allowing for easier and faster creation of new Head Up Games. The platform not only includes the appropriate hardware, but also is bundled with a software API, to allow designers not specifically trained in software engineering to adopt the platform easily. The resulting platform was evaluated using a case study methodology with two Industrial Design master students. The evaluation not only focused on the usability of the platform, but, more importantly, how the platform affected the design process. The process of developing the platform and the evaluation are documented in Chapter 4. The main conclusion of the case studies was that indeed the platform allowed designers to rapidly create mobile games, without the need to spend time and effort on typical embedded hardware/software issues, as battery management or electronic wiring of components. Furthermore, it seems that the platform hardware is suitable to support interaction styles in Head Up Games. However, during the evaluation we came across challenges that have been identified earlier in the End-User Programming domain which limit the ability of designers to create their own games: for example managing software complexity and debugging the software are beyond the reach of non-professional software creators and an environment that will support them needs to provide appropriate scaffolding. These issues require further research.

Finally, based on our experiences using low-fi mockups for the evaluation of Head Up Games, we argue that the design process benefits from using high-fi prototypes from an early stage. In Chapter 3 we describe a study that follows the development of three Head Up Games, designed in a rapid, iterative process. We conclude that by using high-fi prototypes from the start, implications of interaction on the game experience can be addressed earlier and better in the design process. Seemingly small technological issues can have a big effect on the game play that can now be detected within the first week of development. In contrast, in a standard User Centered design process, which increases the fidelity of prototyping only during late phases of the design process, such issues would have been identified too late. Furthermore, in Chapter 3 we reflect on a few other lessons learned on the design process of Head Up Games, such as how to involve children (and adults) in the design process.

CONTRIBUTIONS

From the work in this thesis, several contributions can be identified. The main contribution is the conception, rationalization and demonstration of a new game genre: Head Up Games.

Several minor contributions can be derived from the main contribution, each of the contributions addresses one or more challenges as put forward in Chapter 1:

- we have developed and validated RaPIDO, a platform for prototyping tangible and embodied interaction in outdoor games for children (creation challenges);
- we have proposed an iterative design process, using high-fidelity prototypes early in the design (design challenges);
- we have gained insights into the use of OPOS, an observation scheme for quantifying head up play behaviors (evaluation challenge);
- we have developed a new technique for surveying children's opinions in a group setting (evaluation challenge).

Below we will elaborate on each of the contributions.

A NEW GAME GENRE: HEAD UP GAMES

Before the creation of the Head Up Games concept, pervasive games for children were largely centered on mobile devices with screens. The Head Up Games concept contributes to the field of pervasive gaming, as it is a deviation from this: instead of trying to take screen-based computer games outside, we took traditional outdoor games of children as a basis and enhanced the games using technology that is suitable for supporting the behavior patterns as seen in tradition outdoor games. Examples of these games have been presented throughout this thesis.

RAPIDO: A PROTOTYPING TOOL FOR TANGIBLE AND EMBODIED INTERACTION IN GAMES

The creation and evaluation of RaPIDO provides evidence of the added value of a prototyping tool that abstracts the embedded hardware and software for the designer. In contrast to tools such as Arduinos, RaPIDO requires less embedded engineering skills from designers, and therefore can speed up the development process. We argue that the need for such tools will grow; Moore's law predicts that every 18 months hardware will have shrunk half in size. So, truly ubiquitous computing as envisioned by Weiser (1993) is coming well within reach. However, smaller and more sophisticated hardware requires skilled electrical engineers to create working products (or prototypes). For interaction designers to focus on the interactions and not on the technology, a prototyping tool that offers an integrated solution is best. Evidence for this is presented in Chapters 3 and 4.

ITERATIVE DESIGN PROCESS WITH HIGH-FI PROTOTYPES

We argue that, in the case of Head Up Game development (and possibly for ubiquitous developments too), designers or researchers should deviate from the generally accepted way of designing interactive products, i.e. starting with low-fi , non-interactive prototyping. We argue that low-fi prototyping is not useful to inform the design process because these prototypes do not approximate the novel interaction styles used in Head Up Games, threatening the ecological validity of an evaluation. We presented empirical evidence for this by engaging in an iterative design process, presented in Chapter 3.

INSIGHTS INTO QUANTIFIYING PLAY BEHAVIOR

For quantifying the play behaviors in Head Up Games Bakker et al. (2008) proposed the Outdoor Play Observation Scheme (OPOS), because no existing observation scheme could be found for outdoor play. In Chapter 5 of this thesis we describe three design cases in which we applied OPOS, reflecting on its value for evaluating Head Up Games. The time-investment for OPOS is in its current form quite high, and in Chapter 5 we make suggestions for improvement. OPOS has been applied by Tetteroo et al. (Tetteroo et al., 2011), though they did adapt it to their specific needs. We argue that this indicates that there is a need for an observation scheme such as OPOS, though further research should explore what classes of behaviors are most appropriate to observe, and how to observe these.

GROUPSORTER: NEW TECHNIQUE FOR SURVEYING CHILDREN

Gathering children's opinion can be a challenging task, especially when there are many children that need to be surveyed in a short time. To ease this task, we developed GroupSorter, a group interview technique that allows us to discuss with a group of children, while addressing some of the effects that can occur when children are interviewed together (e.g., children might automatically agree with what their peers are saying, without forming their own opinion). We described the application of GroupSorter in Chapter 5, concluding that GroupSorter indeed can support group interviews as intended.

REFLECTIONS

Having presented the thesis and its contributions, this section reflects on some of the issues left unsettled by this research.

In our introduction of this thesis we described our multidisciplinary approach to research Head Up Games. Reflecting on this approach, we argue that taking different perspectives on the challenge was necessary to research the full breadth of Head Up Games. Contributions we made in one field advanced the research in others; the most notable example is the hardware/software development of a platform for Head Up Games, which we then used to research the design process using high-fi prototypes. Then again, taking perspectives from different research fields can also create tension between the fields, as two of the questions below exemplify.

CAN WE MEASURE THE SUCCESS OF A HEAD UP GAME?

One of the most important questions regarding Head Up Games that is only implicitly, but not explicitly, addressed in this thesis, is when is a game a success? Or more specifically, can we measure the success of a game in such a way that this knowledge can guide or inform the design process?

To answer this question, we first return to our initial aim in Head Up Games:

Head Up Games are outdoor, co-located, multiplayer pervasive games that encourage social interaction, physical activity and support adaptable rules, creating a fun experience

The general approach in HCI is to identify the separate constructs that contribute to the experience and try to measure these in isolation. This is the approach we first adopted: we targeted the separate elements of our vision, i.e. to separately measure social interaction and physical activity. OPOS (Chapter 5) was developed to achieve this. The problem is in translating the values obtained into a measure of success. For example, does more physical activity automatically make a better game?

Furthermore, children's fun can be measured using survey tools, e.g., the Fun Toolkit (Read, 2008). Still, fun can originate from many sources: a game can be fun for a player because it gives him a sense of achievement. Alternatively, a game can be fun because players are engaged in a shared activity with their best friends. Or, a game can be fun, because it offers the right level of engagement to players. And a game can be fun, because it is played at a certain location. Simply quantifying the fun does not give insight into the reasons why it was fun.

In contrast, Game Design literature is not interested in quantifying separate constructs of the player experience. For example, Costikyan (2002) argues there are simply too many factors involved that influence the emergent game experience. Knizia states: "The fun and excitement of playing cannot be calculated in an abstract fashion: it must be experienced" in Salen and Zimmerman (2003, p. 25). These views can be explained if we look at the complexity of game design; Salen and Zimmerman note:

"As a game designer, you are tackling a second-order design problem. The goal of successful game design is meaningful play, but play is something that emerges from the functioning of the rules. As a game designer, you can never directly design play. You can only design the rules that give rise to it. Game designers create experiences, but only indirectly." (Salen and Zimmerman, 2003, p. 168).

As both views (HCI and Game Design) are contrasting the question arises, which one is best answering our initial question: can we measure the success of a game?

We argue that for informing the design process the Game Design approach has the most value; a numerical value of fun, physical activity or social interaction simply cannot convey enough information on how to improve a game design. Furthermore, the time-investment required to apply a tool as OPOS is currently quite high; we experienced in our rapid, iterative design process (Chapter 3) that it was unfeasible to apply OPOS. However, this does not mean that measures like OPOS are automatically disqualified. Embodied interaction techniques are relatively new in the field of outdoor play, and an observation tool like OPOS can help to establish a more detailed model that predicts the effect of novel interaction styles on social interaction and physical activity in the game.

COMPARATIVE STUDIES OR RESEARCH THROUGH DESIGN?

In HCI there is a tendency for favoring comparative studies. For example, in the field of tangible user interfaces for children Fails et al. (2005) compare desktop and physical environments for preschoolers. Similarly, Xie et al. (2008) compared interaction styles: a tangible interface, a physical interface and a graphical interface were created for a jigsaw puzzle to evaluate the enjoyment and engagement of children. Both Fails et al. and Xie et al. remark on the difficulty to properly compare the results. When evaluating *Save the Safe* (Soute et al., 2009) we ourselves tried to compare a design where a game object would be manifested as a virtual entity or with a physical game object. Specifically the key to the safe could either be a virtual one that is perceived by the players using vibration on their vest or a physical one that is represented by a ball that can be passed around, seen, grabbed by the players. We came to the conclusion that in the end the usefulness and validity of such comparisons is very limited.

Save the Safe | p. 36

The question arises whether it is valid to compare a novel (be it tangible or digital) interaction to a traditional interaction. To ensure validity of the comparison, all factors, except for the factor under investigation, must be controlled for and kept equal. Here already lies the first problem: as it is recognized that the context in which interactive products are used has a large influence on the user's use of it, HCI research is moving out from lab-settings into the field (especially so where children are involved). In the field it is much more difficult to control all parameters compared to in the lab. Second, assuming it *is* possible to control all parameters, for a valid comparison of designs, the functionality of the designs must be largely kept equal. This implies that the added design possibilities afforded by a new interaction should be disregarded to keep experimental control. This affects the ecological validity of the novel interaction style, as its novel features are not fully deployed. Alternatively, the experimental control is relaxed in favor of the novel interaction style. However, this then raises the question whether the two designs can still be compared. For example, in *Save the Safe*, we compared a game with a haptic interaction style to a game with a traditional interaction style. Safe the Save was implemented in two versions: in one version children wore belts that would start vibrating when a child had possession of the virtual key; in the other version the key was represented by a (physical) ball. Though it might seem that this is only a minor difference in interaction style, it had a great impact on the game play: directly related to the physicality of the ball is the visibility of the ball. The fact that the ball was visible deeply changed the resulting game experience compared to the haptic interaction style, the latter is invisible for other players, and this brought elements of uncertainty and excitement in the game that were not available in the game with the ball.

Instead of conducting a comparative study to identify the added value of novel interfaces, we found that, at least for Head Up Games a research through design approach (Zimmerman et al., 2007) is more suitable. Taking this approach, we acknowledge that the situations we are designing for are too complex (also referred to as "wicked" problems (Rittel and Webber, 1973)) to tackle in a scientific way and should be approached in a more holistic approach to better take into the multitude of factors that are of influence.

WHERE ARE THE ADAPTABLE RULES?

In our goal, we stated that Head Up Games need to support adaptable rules, so that children can adapt the game to the context they are playing in. So far, the "adaptable rule" goal did not receive much attention in this thesis, though we did engage in two studies that gave insight on the adaptable rules.

Concretely, we distinguish two categories of rules: the first category is rules that are programmed in the game technology. For example, in *Camelot* the time to acquire a resource is programmed in the technology. The second category is rules that are not prescribed by technology, but can be (implicitly) agreed upon by the players. For example, in an early version of *F.A.R.M.* players agreed on the starting distance between players.

Camelot | p. 34

F.A.R.M. | p. 69

With respect to the first category, in a study we explored whether or not children could change the programmed rules (Toering et al., 2010). Our aim was to enable children to adapt some parameters of a game, right before playing the game itself, and more importantly also during playtime. Though a possibility was to let the children set the rules at a computer and upload to the game devices, this would take more time and effort. Also, by requiring a computer interface, one can imagine that the flow of the game activity would be broken if during playtime a rule needed to be changed. Instead, we explored the possibility of setting the game parameters in an outdoor context, i.e. without needing a computer. We argue that this would better support children in just take up the technology, start to play and adapt the rules whenever necessary right on the spot. For

Figure 6.1 Top: Swinxs game console and rule cards. Left: Children setting the rules before playing a game

this study we used a commercially available game platform, Swinxs* (see Figure 6.1) because it carried several of the characteristics of Head Up Games. The rules could be set using small, tangible, game cards that the Swinxs could identify. The result of the study showed that the children indeed were able to understand and use the cards to manipulate game parameters, e.g., setting up the game for a specific number of players, or selecting whether to play a game in teams or compete individually.

* http://www.swinxs.com

With respect to the second category, rules not programmed by the game, we argue that, before children start to *adapt* the rules, they first should *adopt* the game, and that happens when a game is played multiple times. In our studies the games mostly were tested a single time only, and though we did find evidence that children were able to change the rules (see Chapter 3), this was in the setting of developing a game. To test whether children repeatedly play the games and indeed adapt the rules, a longitudinal study is needed, in which children have constant access to the games and can play the games whenever they want. Such a study brings about many challenges. First, as a researcher it is virtually impossible to be present at all times, so careful consideration is needed on how the data must be gathered without the researcher being present. Second, the technology of the game under evaluation should be robust and durable enough to survive long-term testing, which is typically not the case when prototypes are concerned.

To explore these issues we set up an evaluation study using again the Swinxs, because at that time the hardware for our Head Up Games was not yet robust enough. The aim of the study was to explore the possibilities and difficulties of such a longitudinal setup, see Soute and Sturm (2011). More specifically regarding Head Up Games, we were interested to find out if we, without being present, could capture data on the context in which the children played the game and whether they had changed any rules that were not explicitly enforced by the game technology. Such data is not retrievable using for example log data, and we decided to ask the children to shortly interview each other after each game play. The study was carried out in a school setting, running for four weeks. During free play children were allowed, but not required, to play with the Swinxs.

The main conclusion of our study was that indeed the data gathering was a challenge. Though the children repeatedly played with the Swinxs, unfortunately the school schedule allowed for very little time for recording the children's interviews. However, we found that the data gathered through self-interviewing was quite rich in detail. Furthermore, we conducted three sessions with the children in which we first observed the game play and next interviewed the children. Interestingly, we observed that the children indeed seemed to have changed the games, but when asked in the interview whether they had changed any rules they denied doing so. Pointing out to them instances of rule changes we observed in the preceding playtest, it became clear to us that children interpreted "changing the rules" as "cheating", which they did not believe they did, or at least did not want to admit to.

Now that the RaPIDO platform is ready we would be able to explore the possibilities of adapting both categories of rules further. With regard to changing the parameters of game rules, we could further explore the possibilities for using tangible objects to rapidly set parameters, though careful consideration has to be given to how it is reflected to the children what particular rules have been set.

Furthermore, using RaPIDO, we could now run a longitudinal study with Head Up Games. With the insights gained in the study described above, we would make some changes to the setup: a) either change the context to an after school setting, or make better arrangements with the school for allowing time for the self-interviews and b) more carefully phrase our questions, to ensure that children did not misinterpret "changing the rules". The rest of the setup would remain largely similar; combining the self-interviews, observations and group interviews gives a clear picture of how the game play is evolving.

THE FUTURE OF HEAD UP GAMES

The previous section reflecting on this thesis hinted already at interesting directions for future research, such as researching whether Head Up Games are adopted by children in the long run, and whether children are able to adapt the games to their context. Further, each chapter in this thesis has provided directions for future research specifically related to the topic of each chapter. In this section we will present more general directions for future work.

INTERACTION DESIGN IN HEAD UP GAMES

When designing RaPIDO our approach for identifying appropriate interaction styles for Head Up Games has been rather pragmatic: we looked at existing outdoor games and we reviewed the Head Up Games that had been designed until then. Future research could take a more systematic approach in exploring the interaction patterns that exist in Head Up Games (see also Figure 4.1 on p. 98). Each of these patterns affect the game experience in different ways and can be supported in various ways by interaction styles. For example: an abstract interaction pattern in a game is the act of tagging. In a game this could be reflected in the following two ways: a tag can be physical (i.e. not implemented in technology), or alternatively, it can be based on proximity (as was done in *Save the Safe*). Both styles affect the game experience and player behavior in different ways. Future research could map out how interaction styles affect play behavior and experience, which could result in design guidelines for Head Up Games. Furthermore, findings concerning novel interaction styles could inform eventual redesigns of RaPIDO.

EXPLORING THE DESIGN SPACE OF HEAD UP GAMES

Not only the interaction styles are of interest for further research in Head Up Games - the types of games themselves are too. For example, the games we have developed so far often resembled tag and hide-and-seek. The question is whether the Head Up Games concept is limited to these types of games or whether it can support more varied play as well. One approach could be to have various kinds of designers, or more broadly, various people who work with children, design Head Up Games and evaluate the types of games that they come up with. In related work, Márquez Segura et al. (2013) have been inspired by the Head Up Games concept and have created 'body games' - games from which the enjoyment comes from bodily engagement. Furthermore, Mustafin et al. (2012) created a game with a flying, tangible object, with Head Up Games in mind. Both efforts indicate that the Head Up Games concept can inspire a wider range of games than presented so far, though more research is required.

PROTOTYPING TOOLS SUPPORTING AN ITERATIVE DESIGN PROCESS AND EVALUATION

In Chapter 3 and 4 we have seen that the design process of Head Up Games benefits from an iterative approach. However, in the iterative approach we found we had limited time to evaluate the games using the methods proposed in Chapter 5, more specifically OPOS turned out to be too time consuming. Future work could review how measures that are now gathered through observation can be automatically collected. For example, Carter et al. (2008) and also Tang et al. (2011) suggest that for the field of pervasive and ubiquitous computing to advance, attention should be paid to the ecological validity of evaluations. One way to do this is, when designing prototyping tools, to not only design for functionality, but also for evaluation; i.e. a prototyping tool should support designers in gathering and analyzing test data. We reached a similar conclusion in Chapter 5. Such support of a prototyping tool becomes especially useful when longitudinal evaluations are carried out and researchers cannot be present at all times. Prototyping tools that can capture data concerning emergent play behavior are of value then.

Furthermore, to better support designers to explore technology early on in the design process, we argue that the threshold for adopting it should be low. Currently, RaPIDO offers a textual programming interface, but future research could investigate whether a “sketch”-like tool, or a visual programming environment, could help designers implement their ideas faster, and, as a consequence, enable them to more easily explore the design space and engage in an iterative design process.

HEAD UP GAMES IN A BROADER CONTEXT

What is striking of Brueghels painting (see Figure 2.1 on p. 26) is that the games he drew were common for children’s games in the 16th century. Up until recently, these games were still pretty much monopolizing the games played by children. The digital and media revolution changed that: today the entertainment offered to children is more varied; children can choose to play on their computer, their game consoles, watch TV shows and movies, be entertained by a myriad of electronic toys, and, we would almost forget, they could play outside!

Concerns are raised that children spend too much of their time sedentary and inside, resulting in for example obesity, e.g., Rey-López et al. (2008). Furthermore, children spend less time in outdoor, free play (Gleave, 2009), leading Louv (2008) to argue that children these days suffer from a "nature-deficit". Consequently, children may have fewer opportunities to engage with their peers in free play. It is argued in related child development theory, that playing with other children is essential for a child's development; it creates a safe environment to explore different social behaviors, for example conflict, competition and collaboration (Scarlett et al., 2004).

All the above reasons should be motivators to let our children play and also play outside more. The question is: can Head Up Games play a role in this? The answer is yes, though we do not presume that Head Up Games are the exclusive answer.

The children growing up today are more and more becoming digital natives and technology is becoming more and more integrated in everyday life. Though advances in technology allow us to use it on the move, our interactions with it are still predominantly using the visual modality. Smartphones, tablets and portable game technology, e.g., the Nintendo DS, encourage the mobility of electronic play. However, the games played on these devices are largely screen based, often derivatives of successful PC games, and therefore are not designed to encourage social interaction and physical activity. But the trend has been set: it seems that children these days pretty much *expect* the ubiquity of technology.

Head Up Games can combine novel interactions with game play that has appealed to children throughout the ages. This way, Head Up Games can add to the current offerings of entertainment for children, while at the same time encourage children to play outside more.

As a final remark we would point out challenges in a broader context of our society. The availability of electronic toys and computer games is not the only reason children are staying indoors. Research has shown that parents are reluctant to let their children play outdoors, because they find the (city) surroundings too dangerous for a child to play in unsupervised (Veitch et al., 2006). Furthermore, the number of children in one family has decreased, and combining that fact with the observation that children currently have less free time available (Gleave, 2009) it might be a challenge for a child to find a group of children to play with. Taking the above into account we see a future for Head Up Games in an organized, outdoor, context, e.g., for scouts, at schools, at after school care centers, or similar venues for children.

THE FUTURE OF THE RAPIDO PLATFORM

Inevitably, when a piece of technology is produced, it is outdated the moment it is created. The same is true for the RaPIDO platform. However, we argue that the point has been made: there is a need for an integrated, prototyping tool, though there still remains some work to be done. In Chapter 4 we discussed the trade-offs in creating the platform, and one of the main restrictions for not creating smaller, and more appropriate for children, sized devices, was the costs for manufacturing. As prices will continue to decrease, creating a similar yet smaller platform will become feasible.

An obvious question to address is why not use mainstream off the shelf devices like smartphones as an implementation platform? There are several reasons for which we did not pursue this. First, though the technology in current mobile phones is rapidly approximating the technology on the platform, for now it is not sufficient for supporting Head Up Games. Features commonly used in Head Up Games include RFID sensing and distance measurement. Currently, not all mobile phones include RFID technology, though there seems to be a rise in new models that are adopting it. Furthermore, many smart phones have a GPS chip for determining a location. Though this could potentially be used for distance measurement, issues with losing the GPS signal in “shaded” areas and the accuracy and latency of these GPS chips prevent useful application in our games, where children play in close range, and the distance needs to be measured rather accurately and fast.

More important, however, is the design statement we wanted to make, by *not* designing our games on mobile phones. Our aim was to emphasize the aspects of traditional play, i.e. the physical activity and the rich, social interaction. And though, indeed such games *could* have been designed on mobile phones, in related work on outdoor games for children, this did not happen. Most games had a strong focus on the visual aspects and virtual worlds, which we argue interfered with the children's social interaction and physical activity in the real world. Therefore, we found it important to make a firm design statement, which clearly excluded a screen as an interaction modality. Finally, and arguably the most important reason for not using a mobile phone as a platform, is that we see the Head Up Games devices as toys that support play, similar as a ball or a hoop. Therefore, they should look and feel like a toy. By offering dedicated and integrated hardware, designers are free to design their own form around the electronics, creating an interactive game device, specifically tailored for outdoor play.

Concluding, is there a future for the RaPIDO platform? In the short term many more Head Up Games can be developed on it, e.g., to further explore the design space for these game. But in the mid-term, i.e. in a few months to years, its current incarnation will feel antique. Furthermore, most probably by then mobile phones will have evolved to a level that they can support the development of Head Up Games. In that light, we should see the platform as a proof of principle that there is a need for these types of platforms that support interaction designers in designing and deploying outdoor games, and ubiquitous systems in general.

CLOSING REMARK

The world of children today is different than the one we grew up in; advances in technology have had (and still have) a considerable influence on this change. It can be argued that current gaming technology negatively influences children's lifestyle, e.g., by fostering sedentary behavior, or by giving rise to violent behaviours. Furthermore, concerns are raised that displacing face to face interactions with remote ones over social media, possibly skews the social development of children. Rather than passively observing these developments, this thesis aimed to set a new direction for games for children. We claim that recent developments in mobile technology give designers the ability to counter this trend and design novel play experiences for children, combining the appeal of traditional games with that of novel interaction paradigms like pervasive computing and tangible and embodied interaction. In this way, games keep the best of both worlds potentially countering the worries discussed often among scholars and society at large mentioned above.

In Chapter 2 we paraphrased Marzano's maxim in "La Casa Prossima Futura" (Marzano, 2008):

"The games of tomorrow will look more like the games of yesterday than the games of today"

The work presented in this thesis signifies a first step of the creation of a novel genre of children's games. It will obviously take more work and effort to truly insert these types of games into children's everyday lives. Still, the direction has been set and so we venture to say: tomorrow starts today.

REFERENCES

Bakker, S., Markopoulos, P., and Kort, Y. de. (2008). OPOS: an observation scheme for evaluating head-up play. In Proc. of NordiCHI '08 (pp. 33–42). ACM, Lund, Sweden.

Carter, S., Mankoff, J., Klemmer, S. R., & Matthews, T. (2008). Exiting the Cleanroom: On Ecological Validity and Ubiquitous Computing. Human–Computer Interaction, 23(1), 47–99. doi:10.1080/07370020701851086

Fails, J. A., Druin, A., Guha, M. L., Chipman, G., Simms, S., and Churaman, W. (2005). Child's play: a comparison of desktop and physical interactive environments. In Proc. of the 2005 Conf. on Interaction Design and Children (pp. 48–55). ACM, Boulder, Colorado.

Gleave, J. (2009). Children's time to play: A literature review. London: Play England. Available online at: http:*www. playday. org. uk/pdf/Childrens-time-to-play-a-literature-review. pdf (Accessed Nov. 2011).*

Ko, A. J., Myers, B. A., & Aung, H. H. (2004). Six Learning Barriers in End-User Programming Systems. In Visual Languages and Human Centric Computing, 2004 IEEE Symposium on (pp. 199 –206). doi:10.1109/VLHCC.2004.47

Louv, R. (2008). Last child in the woods: Saving our children from nature-deficit disorder. Algonquin Books.

Marzano, S. (2008). Ambient Culture. In J. L. Encarnação and E. H. L. Aarts (Eds.), True Visions: The Emergence of Ambient Intelligence (1st ed. 2006. 2nd printing.). Springer.

Márquez Segura, E., Waern, A., Moen, J., & Johansson, C. (2013). The design space of body games: technological, physical, and social design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 3365–3374). New York, NY, USA: ACM. doi:10.1145/2470654.2466461

Mustafin, R., Wehner, J., Sattler, W., & Gohlke, K. (2012). T.F.O.: tangible flying objects. In Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (pp. 193–196). New York, NY, USA: ACM. doi:10.1145/2148131.2148173

Read, J. C. (2008). Validating the Fun Toolkit: an instrument for measuring children’s opinions of technology. Cogn. Technol. Work, 10(2), 119–128.

Rey-López, J. P., Vicente-Rodríguez, G., Biosca, M., and Moreno, L. A. (2008). Sedentary behaviour and obesity development in children and adolescents. Nutrition, Metabolism and Cardiovascular Diseases, 18(3), 242–251.

Rittel, H. W. J., and Webber, M. M. (1973). Dilemmas in a general theory of planning. Policy Sciences, 4(2), 155–169.

Salen, K., and Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals (illustrated edition.). The MIT Press.

Scarlett, W. G., Naudeau, S. C., Salonius-Pasternak, D., and Ponte, I. C. (2004). Children's Play (1st ed.). Sage Publications, Inc.

Soute, I., Kaptein, M., and Markopoulos, P. (2009). Evaluating outdoor play for children: virtual vs. tangible game objects in pervasive games. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 250–253). ACM, Como, Italy.

Tang, L., Yu, Z., Zhou, X., Wang, H., & Becker, C. (2011). Supporting rapid design and evaluation of pervasive applications: challenges and solutions. Personal and Ubiquitous Computing, 15(3), 253–269. doi:10.1007/s00779-010-0332-6

Tetteroo, D., Reidsma, D., Van Dijk, E., Nijholt, A., Camurri, A., Costa, C., and Volpe, G. (2011). Design of an interactive playground based on traditional children's play. In Proc. of the 4th Int. Conf. on Intelligent Technologies for Interactive Entertainment (INTETAIN 2011) (Vol. 78, pp. 129–138).

Toering, E., Soute, I., and Markopoulos, P. (2010). Rule customization in head-up games. In Proc. of the 3rd Int. Conf. on Fun and Games (pp. 144–148).

Veitch, J., Bagley, S., Ball, K., and Salmon, J. (2006). Where do children usually play? A qualitative study of parents' perceptions of influences on children's active free-play. Health & Place, 12(4), 383–393.

Weiser, M. (1993). Some computer science issues in ubiquitous computing. Commun. ACM, 36(7), 75–84.

Xie, L., Antle, A. N., and Motamedi, N. (2008). Are tangibles more fun?: comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. In Proc. of TEI '08 (pp. 191–198). ACM, Bonn, Germany.

Zimmerman, J., Forlizzi, J., and Evenson, S. (2007). Research through design as a method for interaction design research in HCI. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (pp. 493–502). ACM, San Jose, California, USA.

Appendix A

Contributions

Most of the work presented in this thesis has been carried out in collaboration with other researchers and students. Most notably some of the Head Up Games referred to in this thesis have been designed and/or implemented by students under supervision of either myself or Panos Markopoulos.

In this section I will highlight for each main chapter what my personal contribution has been. The publications that each chapter is based on are largely mine, with comments and feedback of the other authors.

CHAPTER 2 - INTRODUCING HEAD UP GAMES

based on:

> Soute, I., Markopoulos, P., & Magielse, R. (2009). Head Up Games: combining the best of both worlds by merging traditional and digital play. Personal and Ubiquitous Computing, 14(5), 435–444. doi:10.1007/s00779-009-0265-0

The concept of Head Up Games was formed by myself and Panos Markopoulos. The chapter presents several examples of Head Up Games: *Camelot*, *Save the Safe* and *HeartBeat*.

Camelot has been created before the start of this PhD thesis by myself, Janneke Verhaegh and Angelique Kessels; the work was done in equal parts. The concept of *Save the Safe* was largely based on the game concept of *Stop the Bomb*. However, the implementation of *Save the Safe*, and the evaluation and analyses of the results were mine. *HeartBeat* was created, implemented and evaluated by Remco Magielse.

CHAPTER 3 - DESIGNING HEAD UP GAMES

based on:

> Soute, I., & Markopoulos, P. Designing for Embodiment in Outdoor Interactive Games for Children. In A. Nijholt (Ed.), Playful User Interfaces: Interfaces that Invite Social and Physical Interaction. Springer. To appear.

> Soute, I., Lagerstrom, S., & Markopoulos, P. (2013). Rapid Prototyping of Outdoor Games for Children in an Iterative Design Process. In Proc. of the 12th Int. Conf. on Interaction Design and Children. New York: ACM. To appear.

The first part of this chapter mostly reflects on experiences I gained during the game design and evaluation of Head Up Games.

The second part of chapter 3 relating to using prototypes describes a study which I have run together with Susanne Lagerstrom, a master student under my supervision. The games *Invade the Castle* and *Follow the Light* were designed by Susanne and implemented with my help. *Save the Safe* and *F.A.R.M.* were by my design and implementation. Susanne did most of the work for preparing the study, the sessions were largely run by me. Analysis was done by both me and Susanne.

CHAPTER 4 - CREATING HEAD UP GAMES WITH RAPIDO

based on:

> Soute, I., & Markopoulous, P. Creating Head Up Games with RaPIDO: a platform for rapid prototyping of mobile outdoor games for children. Submitted to ACM Transactions on Computer-Human Interaction (TOCHI)

The conception, creation and evaluation of the RaPIDO platform have all been my work. For the hardware implementation Herman Aartsen offered technical support: he designed and implemented the electronic layout. For the creation of the casing I had support of Chet Bangaru, who created the 3D models of the devices.

CHAPTER 5 - EVALUATING PLAYER EXPERIENCE IN HEAD UP GAMES

based on:

> Soute, I., Bakker, S., Magielse, R., & Markopoulos, P. (2013). Evaluating player experience for children's outdoor pervasive games. Entertainment Computing, 4(1), 25–38. doi:10.1016/j.entcom.2012.09.003

The Outdoor Play Observation Scheme (OPOS) was created by Saskia Bakker, supervised by Panos Markopoulos. GroupSorter was created by myself. The studies in which both OPOS and GroupSorter were applied were run and analyzed by Saskia Bakker (*LightHouse*), Remco Magielse (*HeartBeat*) and myself (*Save the Safe*) respectively.

Appendix B

RaPIDO Manual

INTRODUCTION

Welcome! This manual will help you get started using the HUGs prototypes and developing your own games or prototype mobile interactions.

If you have any questions, please feel free to contact me at: i.a.c.soute@tue.nl

THE SOFTWARE

SOFTWARE INSTALLATION

Arduino

To be able to build sketches in Arduino, please install Arduino v022 from the Arduino website. For now, the newest version of Arduino is not supported.

After installation, copy the folder `HUGsSW` into the `libraries` folder in your sketches folder (to find the location of your sketches folder, check the Arduino preferences). If the `libraries` folder does not exist, create one first and then copy the `HUGsSW` folder into it. Finally, locate the existing file `boards.txt` in your Arduino application folder. Replace that file with the `boards.txt` file located in the `HUGsSW` folder, so the Arduino environment knows how to upload to the HUGs boards. If the Arduino application is running, close it and reopen it.

Now, you can start a new sketch for the HUGs boards! Don't forget to select the right board `Tools -> Board -> HUGS 8Mhz` before uploading.

Note: many of the examples in this manual start with `#include <WProgram.h>`. This line of code is only needed when programming the boards from Eclipse (see next section). You can safely ignore this line when programming in the Arduino environment.

Eclipse

Before you can start using the HUGs platform you will have to install a few tools on your computer, namely:

- Eclipse IDE for C/C++
- Arduino. Please get version 022 as the newest version of Arduino is not yet supported. Also install the FTDIUSBSerialDriver that comes with Arduino.
- Arduino plugin for Eclipse. Here you can find more information on how to install it.

Install each tool with the default options. Both Eclipse and Arduino do not need further setup. However, the plugin requires some steps to set up. Instructions can be found [[add link]] (Mac OS X only, however they will not vary much on another platform).

Finally, in the Arduino directory (Windows) or package (Mac OS X) a file called `boards.txt` exists. This file needs to be replaced by the new $boards.txt$ file, as this new file includes the definitions of the HUGs devices.

To connect and get feedback from the board through USB you can use any serial monitor, for example goSerial (Mac OS X).

To add the HUGSimpl and HUGSapi source files to your eclipse workspace, do the following:

Go to the properties of your project -> C/C++ general -> Paths and Symbols and add the folders to both the include tab and the source location tab. See Figure A.1 and Figure A.2.

DESIGN OVERVIEW

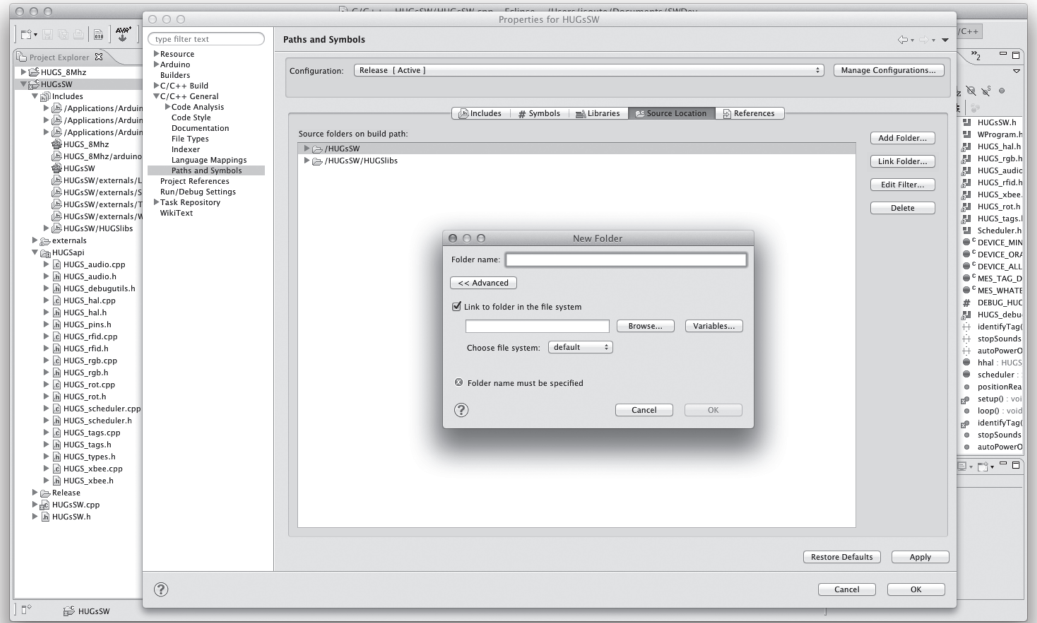


Figure A.1 Adding source folders to the project

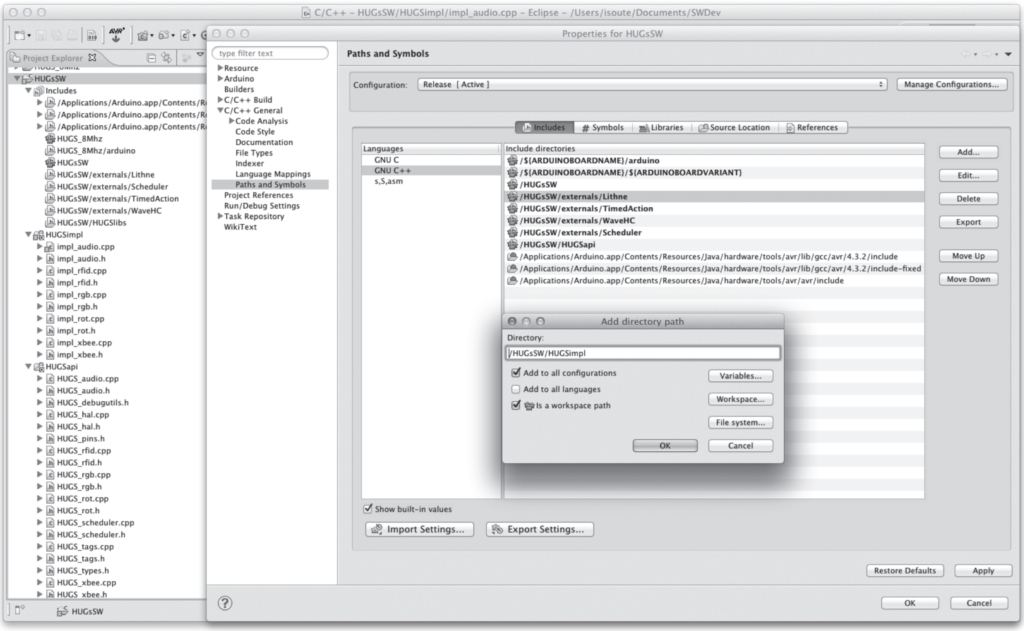


Figure A.2 Adding include folders to the project

Figure A.3 shows an overview of how the software is designed. As you can see the software has been divided into several parts. First, on the right, dedicated libraries have been written, that interface directly with the hardware. Second, an coordinating component called HAL (hardware abstraction layer) has been created. Basically, this component 'hides' the dedicated libraries from the rest of the software. This ensures that, if a dedicated library might change in the future because of changes in the hardware, this does not directly affect the rest of the software. From your point of view, you need not worry about this part of the software. To program your games, you will use the API (application programming interface). The API offers functionality that you most probably need for programming your games.

In the next section a few general concepts regarding the Arduino software will be further detailed. Subsequently in the next sections each component and the functionality it offers will be explained with examples.

ARDUINO

As explained in section - The hardware the HUGs devices run on an adapted Arduino ATMega processor. Though programming for Arduino can be done in the dedicated Arduino IDE (integrated development environment), this manual assumes you will be programming your software using the Eclipse IDE. If you are not very familiar with programming in C/C++, I suggest you start out with the example files and go from there, instead of starting from scratch.

Basically, any Arduino program exist of two functions: `setup()` and `loop()`. At the start

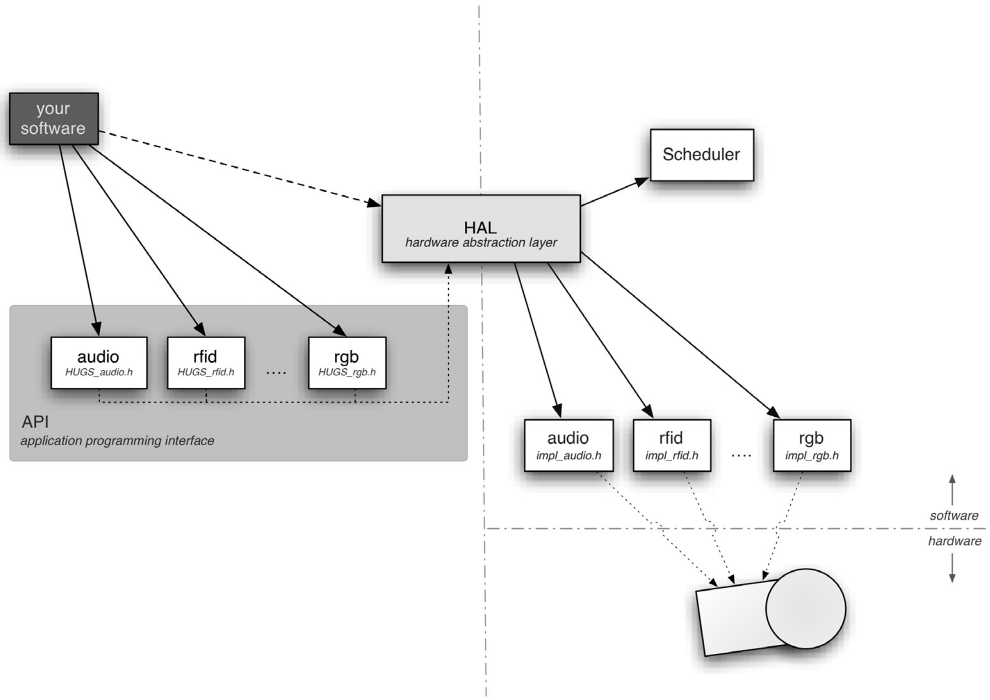


Figure A.3 Overview of the software

of the program `setup()` is executed once and then the `loop()` will be executed indefinitely. Before you start programming your game, think well on how the software should be designed as to accommodate this.

API

In this section the basics of each hardware component and it's corresponding software module will be explained. Often modules will offer more functionality than described here, to find out check the Header file (**.h) file. There you will also find comments with each function that describe what it does.

VIBRATION MOTOR

The code for operating the vibration motor is pretty straight forward: it can be switched on, and it can be switched off. As it is so simple, its code is not in a separate file, but has been located with HUGS_hugs.h. For an example, see below:

```
include <WProgram.h>
include <HUGS_hugs.h>
//The setup function is called once at startup of the sketch
void setup()
{
   hugs::initialize();
}
// The loop function is commentsalled in an endless loop
void loop()
{
   // Perform basic functionality (ALWAYS do this!)
   hugs::update();
   // Switch in
   vibra::on();
   hugs::delayGame(1000);
   vibra::off();
   hugs::delayGame(1000);
}
```

For your convenience, one extra function has been created. You can set a duration to indicate how long the vibration motor should be switched on, like so:

```
vibra::on(500);
```

In this example, the vibration motor will automatically switch off after 500 ms (note that the duration is defined in milliseconds).

HAL AND SCHEDULER

The hugs module is the "master"-module of the software. You always need this, as it takes care of some important maintenance routines for you. For the HUGs devices to work some basic processes need to be run every once in a while. To ensure this, you must call the two functions of `hugs:: initialize()` and `hugs::update()` in `setup()` and `loop()` respectively. Here is shown how:

```
#include <WProgram.h>
#include <HUGS_hugs.h>
void setup()
{
       hugs::initialize();
       // Add your code here.
}
void loop()
{
       hugs::update();
       // Add your code here.
}
```

`hugs::initialize()` initializes the HUGs board for you. It sets all in and out pins to the components correctly and puts all components in a low energy state, to make sure that they do not draw unnecessary battery power. To power up the components that you want to use, use the `enable()` function of that particular component.

`hugs::update()` performs necessary routines e.g., checking the battery level and switching off the device once the battery level drops below a minimum level. It also checks the whether the user presses the on/off button (if the button is pressed for three seconds the power will switch off).

Furthermore hugs offers means for scheduling tasks. For example, you want something to happen repeatedly, e.g., every 5 seconds. Simply write a function that does the things that you want done, and then offer that function to HAL for scheduling, like so:

```
#include <WProgram.h>
#include <HUGS_hugs.h>

void doSomething();                    //note: declaration for-
warding
void setup()
{
       hugs::initialize();
       hugs::startRepeatingTask(5, doSomething);
       // Add your code here.
}
void loop()
{
       hugs::update();
       // Add your code here.
}
void doSomething()
{
       // Add your code here, I.e. the code that you want to be
executed repeatedly
```

```
}
```

The call `hugs::startRepeatingTask(5, doSomething);` schedules the execution of `doSomething()` for every 5 seconds. Note that you need to consider *when* you call this function. In the example above it is called within `setup()`, thus it will only be called once. If you would call it in `loop()`, it would result in a new task being scheduled *every time* `loop()` is executed!

If necessary, you can stop a repeating task too. startRepeatingTask returns an integer value. This is the 'identification number' of the scheduled task. To stop a task, you use stopRepeatingTask, and set the ID as an argument, e.g.:

```
include <WProgram.h>
include <HUGS_hugs.h>
include <HUGS_cap.h>
int idDoSomething;
void doSomething();
//The setup function is called once at startup of the sketch
void setup()
{
   // Initialize basic HUGs functionality
   // ALWAYS call this function!
   hugs::initialize();
   hugs::autoPowerOff(2);

   idDoSomething = hugs::startRepeatingTask(5, doSomething);
   // Enable the cap module.
   cap::enable();
}
// The loop function is called in an endless loop
void loop()
{
   // Perform basic functionality at the start of a new loop (AL-
WAYS do this!)
   hugs::update();

   if (cap::swipeStarted())
   {
          // The start of the swipe is detected
          hugs::stopRepeatingTask(id);
   }
}
```

Finally, if you need to schedule a task in the future for one-time execution, i.e. not repeating, you can use:

```
hugs::scheduleTask(5, doSomething);
```

This will execute doSomething() *once*, after 5 seconds.

Overview of all hugs functions

Here's the list of all functions that the hugs module offers. Check the HUGS_hugs.h header file for detailed information on each function.

```
void initialize();
void update();
void powerOff();
void autoPowerOff(unsigned int min);
void delayGame(unsigned long duration);
int startRepeatingTask(unsigned int sec, void (*func)());
void stopRepeatingTask(int ticket);
void scheduleTask(unsigned int sec, void (*func)());
```

AUDIO

The audio component provides several functions for the playback of WAV files that are on the SD card. Before the audio can be used, you must call it's `initialize()` function in `setup()`. See below:

```
include <WProgram.h>
include <HUGS_hugs.h>
include <HUGS_audio.h>
//The setup function is called once at startup of the sketch
void setup()
{
   hugs::initialize();
   audio::enable();
   audio::setVolume(3); //setting the audio volume is optional
}
// The loop function is called in an endless loop
void loop()
{
   // Perform basic functionality (ALWAYS do this!)
   hugs::update();
   audio::playComplete("DOG.WAV");
}
```

Note that if you don't use `setVolume()` the volume will be at its maximum level. The argument of `setVolume()` is an integer; the higher the value, the lower the volume.

Other functions the audio component offer include:

`void playComplete(char *name)`

Completely plays the file called 'name'. This is a blocking function.

`void startPlayFile(char *name)`

Start playing the file 'name'. Immediately returns control to the calling function, without waiting for the file to finish playing.

```
void startPlayRepeat(char *name)
```

Same as above, however once the end of the file is reached playback will restart automatically. Can be stopped by calling the function stop()

```
void stop()
```

Stops playback if a file is currently being played (both in normal mode, as well as in repeating mode)

CAPACITIVE SENSORS

On the bottom of the device there are two "bobbly" areas. These areas are touch sensors. See Figure A.4.

The most easy way to use these sensors is by checking whether or not a 'touch' is detected:

```
bool isSensor1Touched = cap::touchSensor1();
bool isSensor2Touched = cap::touchSensor2();
```

Both function return a boolean: true if a touch is detected, and false when nothing is detected. A whole program would look like this:

```
include <WProgram.h>
include <HUGS_hugs.h>
include <HUGS_cap.h>
//The setup function is called once at startup of the sketch
void setup()
{
   // Initialize basic HUGs functionality
   // ALWAYS call this function!
   hugs::initialize();
   hugs::autoPowerOff(2);
   // Enable the cap module.
   cap::enable();
}
// The loop function is called in an endless loop
```

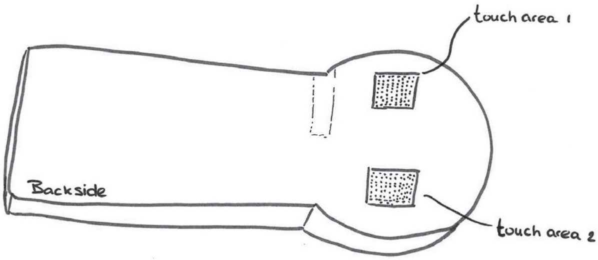


Figure A.4 Location of the touch areas

```
void loop()
{
   // Perform basic functionality at the start of a new loop (AL-
WAYS do this!)
   hugs::update();
   bool isSensor1Touched = cap::touchSensor1();
   bool isSensor2Touched = cap::touchSensor2();
   // Add your code here
}
```

Like every other HUGs component, the capacity module needs to be enabled in the setup() routine.

Besides simply reading the current status of the capacitive sensors, the capacitive module offers also a bit more complex functionality: detecting a swipe movement. This can be implemented as follows:

```
include <WProgram.h>
include <HUGS_hugs.h>
include <HUGS_cap.h>
//The setup function is called once at startup of the sketch
void setup()
{
   // Initialize basic HUGs functionality
   // ALWAYS call this function!
   hugs::initialize();
   hugs::autoPowerOff(2);
   // Enable the cap module.
   cap::enable();
}
// The loop function is called in an endless loop
void loop()
{
   // Perform basic functionality at the start of a new loop (AL-
WAYS do this!)
   hugs::update();
   if (cap::swipeStarted())
   {
         // The start of the swipe is detected
         // Add your code here, e.g. switch on a led
   }
   else if(cap::swipeAborted())
   {
         // Swipe was aborted
         // Add your code here, e.g. switch off a led
   }
   else if (cap::swipeCompleted())
   {
         // Swipe was completed
         // Add your code here, e.g. play a sound file
```

```
    }
}
```

Detecting a swipe is divided into three states: a swipe has been started, a swipe was started but aborted before properly functioning; a swipe moment was completed. As you can see from the code above, these three states can be checked in your `loop()` function, and when either of these states is detected (they are mutually exclusive), a 'true' is returned. The actual swipe can only be performed in one direction. First, the device should be held in the right hand (Figure A.5), with the sensors leaning on the index finger of the left hand. Then, when the left hand is moved to the left (or device is moved to the right), a swipe will be detected (Figure A.6).

It is important to know how the capacitive module works 'behind the scenes', as it might influence your program as it gets complex and demands a lot of the processing power of the processor.

The capacitive module is implemented as a polling mechanism, i.e. the software reads the

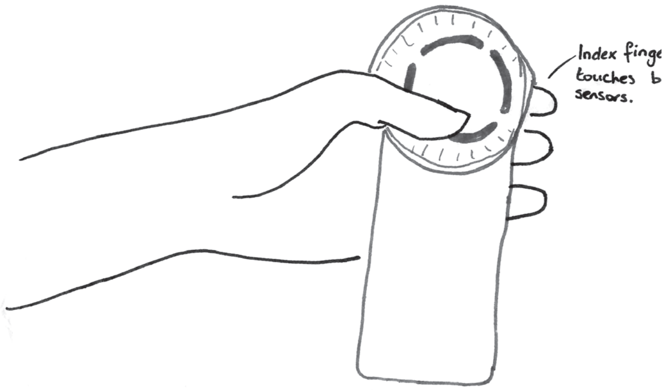


Figure A.5 Start position swipe

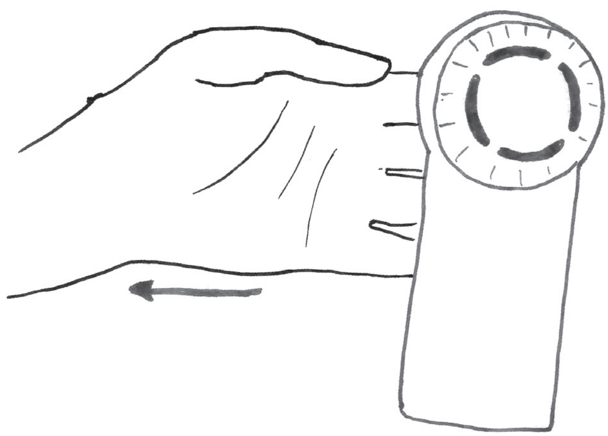


Figure A.6 Swipe movement direction

values of the capacitive sensors every so many seconds. The moment `cap::enable()` is called, a repetitive task is started with a high update frequency (every 50 ms). The update frequency is chosen such that touches are always detected.

However, when you create a program that uses a lot of processing power it might be the case that the polling mechanism of the capacitive module interferes with the workings of your own program. It this is the case either disable the capacitive sensor, or alternatively change the update frequency.

Overview of all cap function

Check HUGS_cap.h for detailed information of each function.

```
void enable();
void disable();
bool touchSensor1();
bool touchSensor2();
bool swipeCompleted();
bool swipeStarted();
bool swipeAborted();
```

RFID READER AND TAGS

The HUGs device can detect RFID tags when they are sufficiently close to the reader. By default the RFID reader is powered down, because it uses a lot of power. To start using it, simply call its `enable()` function. After that you can start using it, by calling its `read()` function.

```
include <WProgram.h>
include <HUGS_hugs.h>
include <HUGS_rfid.h>
include <HUGS_tags.h>
const HUGs_Tag cTAG_B1 = HUGs_Tag(0x21, 0x00, 0xD6, 0x0D, 0x07);
const HUGs_Tag cTAG_B3 = HUGs_Tag(0x21, 0x00, 0xD8, 0x7C, 0xFF);
//The setup function is called once at startup of the sketch
void setup()
{
   hugs::initialize();
   rfid::enable();
}
// The loop function is called in an endless loop
void loop()
{
   // Perform basic functionality (ALWAYS do this!)
   hugs::update();
   //see if there is a tag near
   HUGs_Tag theTag = rfid::read();
   if (theTag == cTAG_B3)
   {
         //Add your code here
   }
}
```

Beginners C/C++ tip it is common practice in C/C++ to make variables that do not change const (as is done in the example above: tag ids are unique and therefore will not change!). Your program will work when you do not use the const qualifier; however, when you do, the compiler will perform an extra check to see if the value of your const variables are not accidentally re-defined somewhere else in the code.

Overview of all rfid functions

Check HUGS_rfid.h for detailed information of each function

```
void enable();
void disable();
HUGs_Tag read(bool doSerialPrint = false);
```

Check HUGS_tags.h for detailed information of the tags functions.

RGB

The HUGs devices have 4 RGB leds, that can light in any RGB color you want. Defining these colors is done in a similar way as defining the tags and a few colors have been defined already for you in HUGS_rgb.h (check the header file for more colors).

From HUGS_rgb.h:

```
const HUGs_RGB RGB_RED = HUGs_RGB(255,0,0); /**< red */
const HUGs_RGB RGB_BLUE = HUGs_RGB(0,0,255);/**< blue*/
const HUGs_RGB RGB_GREEN = HUGs_RGB(0,255,0);
```

If you want to define more colors you can do so in your own header or source file, using the same syntax as above.

The rgb module offers a lot of different functions. Here is a simple example of blinking all the leds:

```
include <WProgram.h>
include <HUGS_hugs.h>
include <HUGS_rgb.h>
//The setup function is called once at startup of the sketch
void setup()
{
   hugs::initialize();
   rgb::enable();

}
// The loop function is called in an endless loop
void loop()
{
   // Perform basic functionality (ALWAYS do this!)
   hugs::update();
```

```
        rgb::allOn(RGB_ORANGE);

        hugs::delayGame(1000);

        rgb::allOff();

        hugs::delayGame(1000);
    }
```

Tip: the rgb module offers the function `blink()` that will accomplish the same as the example above, but with only one line of code :)

One important thing to remember is that, though you can switch the leds on and off independently it is not possible to have two leds switched on using different colors. So, if you set a color this color is applied to all leds until you set another color.

RGB LEDS

To switch on leds individually you use the identifiers `RGB1`, `RGB2`, `RGB3` and `RGB4` to indicate the correct led (see Figure A.7). For example:

```
        rgb::on(RGB1);
```

will switch on the top led. See below for a list of all functionality offered by this module.

Overview of all rgb functionality

Check HUGS_rgb.h for a detailed description of each function

```
    void enable();
    void disable();
    void setColor(HUGs_RGB theColor);
    void setIntensity(HUGs_RGB_ID, unsigned int);
    void toggle(HUGs_RGB_ID);
    void on(HUGs_RGB_ID);
    void on(HUGs_RGB_ID, HUGs_RGB);
    void on(HUGs_RGB_ID theLed, HUGs_RGB theColor,
    int intensity);
    void off(HUGs_RGB_ID);
    void allOn();
    void allOn(HUGs_RGB theColor);
    void allOn(HUGs_RGB theColor, int intensity);
    void allOff();
    void fadeOut(HUGs_RGB_ID);
    void allFadeOut();
    void blink(HUGs_RGB_ID theLed, unsigned long
    duration);
    void startLoop(unsigned long interval);
    void startLoop(HUGs_RGB theColor);
    void startLoop(HUGs_RGB theColor, unsigned long interval);
    void stopLoop();
```

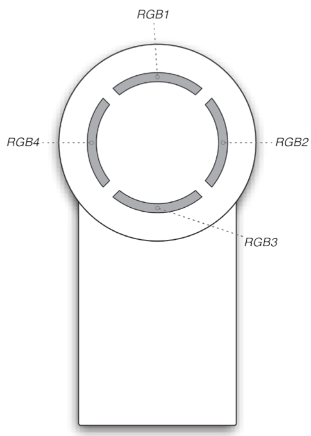


Figure A.7 Location of RGB leds

ROTATION ENCODER (TURNING WHEEL)

The big wheel on top of the device is designed in such a way that it can turn both sides indefinitely. The encoder returns an absolute position of the wheel, an integer between 0 and 85 (approximately), which corresponds with a full turn of the wheel.

The functionality that HUGS_rot.h offers includes simply returning the current position of the wheel, or detecting a turn in clockwise or counterclockwise direction.

The readPosition() function simply returns the current position of the wheel and you can do with this value however suits you in your software.

However, most of the time you will probably do not want to monitor the turning of the wheel yourself, but will you be interested when ever e.g. a quarter (90 degrees) or half turn (180 degrees) has been made. To this end, the rot module offers the functions `detectQuarterTurnCW()`, `detectQuarterTurnCCW()` and `detectHalfTurn()`. Each of these functions accepts a pointer to a function as an argument; that function will be executed once the quarter or half turn has been detected (this is similar to the way the scheduler operates). An example:

```
include <WProgram.h>
include <HUGS_hugs.h>
include <HUGS_rot.h>
void doSomethingAtHalfTurn()
{
   //Add your code here
}
//The setup function is called once at startup of the sketch
void setup()
{
   // Initialize basic HUGs functionality
   // ALWAYS call this function!
   hugs::initialize();
   rot::enable();

   rot::detectHalfTurn(doSomethingAtHalfTurn);
}
// The loop function is called in an endless loop
void loop()
{
   // Perform basic functionality at the start of a new loop
(ALWAYS do this!)
   hugs::update();
}
```

Overview of all rot functionality

Check HUGS_rot.h for a detailed overview of each function

```
void enable();
void disable();
```

```
int readPosition();
int readQuadrant();
void turnWithLed();
void stopTurnWithLed();
void detectQuarterTurnCW(void (*func)());
void detectQuarterTurnCCW(void (*func)());
void detectHalfTurn(void (*func)());
```

RADIO

The radio (XBee chip) can be used for sending and receiving messages to other devices. The use of the radio is quite simple: you can send messages and receive messages.

Sending messages

Before you can use the radio to send (or broadcast) a message a few things need to be done.

First, you must enable the radio in your `setup()` routine (`radio::enable()`). Important to know is that it takes some time for the radio to start up (approx. 10 seconds); after it has started the `setup()` routine will continue.

Second, before messages can be sent, they must be defined first. This can be done in two ways:

1) by assigning a unique ID (an integer)

2) by assigning a unique name (a string)

Once assigned the messaging system works identical. Whichever way you define the messages is up to your own preference - however, always keep in mind that the ID or name has to be unique! If two messages have the same ID/name, this will seriously mess up your run-time code!

The example below shows how you can define a message:

```
//Register the message types, note that MES_UNDEFINED already has
been registered in HUGS_radio.h
const HUGs_Message MES_TAG_DETECTED = 0x01;
const HUGs_Message MES_WHATEVER          = 0x02;
const HUGs_Message MES_TEA_TIME          = 0x03;
const HUGs_Message MES_HAPPY_BIRTHDAY    = 0x04;
```

Note that in de above example, the messages are defined in hexadecimals. Of course you are free to use normal decimal numbers too. What names you give to your messages is entirely up to you - you can choose whatever makes most sense to you in the context of your application.

Alternatively, you can use strings to define messages, like so:

```
//Register the message types, note that MES_UNDEFINED already has
```

```
been registered in HUGS_radio.h
const HUGs_Message MES_COFEE_READY        = radio::registerNewMess
age("Coffee ready");
const HUGs_Message MES_TAKE_A_BREAK       =
radio::registerNewMessage("Take a break");
```

Now the messages have been defined, they can be sent. There are four functions that you can use:

```
void sendToAll(HUGs_Message mes, const int arg = 0);
```

Send a message to all devices. Optionally, you can provide an argument to send along.

```
void sendToTeam(HUGs_TeamID teamID, HUGs_Message mes, const int arg
= 0);
```

Send a message to all devices of a specific team. Optionally, you can provide an argument.

```
void sendToNearestDevice(HUGs_Message mes, const int arg = 0);
```

Send a message to the device nearest to you. A scan will be performed to find out which device is closest, next the message will be sent to that device. Optionally, you can provide an argument.

```
void sendToDevice(HUGs_device theDevice, HUGs_Message mes, const
int arg = 0);
```

Sends a specific message to a specific device. Make sure you have registered the message before sending it. Optionally, you can provide an argument.

The first three functions can be used without explicit knowledge about the other devices in the game. However, if you want to direct a message to a specific device, you will first need to know the address of that device. The address is a 64-bit value, and is defined in two parts: the most significant bit (msb) and the least significant bit (lsb). To find out the address of a device, connect it to your computer, open a terminal and start the device (make sure that the radio is enabled in the code). At startup the address of the device will be displayed in the terminal.

The code below shows how to define specific devices in your code:

```
//Register the device addresses for xbee communication.This is only
necessary
//when you want to send messages to specific devices.
const HUGs_device DEVICE_ORANGE = HUGs_device(0x0013A200,
0x406AB9E7);
```

Receiving messages

Once the radio is enabled, it will continuously scan for incoming messages. Messages that arrive a placed in a buffer until you request them from your code, like this:

```
   HUGs_Message receivedMessage = radio::receive();     //check if
a message was received

   if(receivedMessage != MES_UNDEFINED)
   {
          //received a message! Check which one it is, and act ac-
cordingly
          if (receivedMessage == MES_TAKE_A_BREAK)
          {

                 //add your code here
          }
   }
```

Some messages contain arguments. After you received a message that contains an argument, you can retrieve it using `radio::getArgument()`. For example:

```
   HUGs_Message receivedMessage = radio::receive();     //check if
a message was received

   if(receivedMessage != MES_UNDEFINED)
   {
          //received a message! Check which one it is, and act ac-
cordingly
          if (receivedMessage == MES_TAKE_A_BREAK)
          {
                 int howLong = radio::getArgument();

                 //add your code here
          }
   }
```

Furthermore, you can also get a value of the distance between your device and the one that sent the message, by using `radio::getDistance()`. The lower the value, the smaller the distance is.

Example

Finally, here is a code example of all of the above:

```
include <HUGS_hugs.h>
include <HUGS_rgb.h>
include <HUGS_radio.h>
include <Logging.h>
define LOGLEVEL LOG_LEVEL_INFO
//Register the device addresses for xbee communication.This is only
```

```
necessary
//when you want to send messages to specific devices.
const HUGs_device DEVICE_ORANGE = HUGs_device(0x0013A200,
0x406AB9E7);
const HUGs_device DEVICE_MINT = HUGs_device(0x0013A200,
0x4063E2DF);
const HUGs_device DEVICE_GREEN = HUGs_device(0x0013A200,
0x4063E2CF);
const HUGs_TeamID TEAM_BLUE = HUGs_device::registerNewTeamID("Team
blue");
const HUGs_TeamID TEAM_RED    = HUGs_
device::registerNewTeamID("Team red");
//Register the message types, note that MES_UNDEFINED already has
been registered in HUGS_radio.h
//const HUGs_Message MES_UNDEFINED         = 0x00;
const HUGs_Message MES_COFEE_READY         = radio::registerNewMess
age("Coffee ready");
const HUGs_Message MES_TAKE_A_BREAK        =
radio::registerNewMessage("Take a break");
String myTeam;
//The setup function is called once at startup of the sketch
void setup()
{
   Log.Init(LOG_LEVEL_DEBUG);
   Log.Info("------------------");
   Log.Info("HUGS Device starting up");
   Log.Info("------------------");
   hugs::initialize();
   hugs::autoPowerOff(1);
   // Turn on one led to indicate we're starting up.
   rgb::allOn(RGB_RED);
   //Note: this may take up to 10 seconds.
   radio::enable();
   //wait for team assignment, is saved in HUGs_device
   HUGs_device::assignTeam(cTAG_B8, TEAM_RED, RGB_RED, cTAG_B9,
TEAM_BLUE, RGB_BLUE);
   //All systems GO, switch light off.
   rgb::allFadeOut();
}
// The loop function is called in an endless loop
void loop()
{
   // Perform basic functionality (ALWAYS do this!)
   hugs::update();
   HUGs_Message receivedMessage = radio::receive();     //check if
a message was received

   if(receivedMessage != MES_UNDEFINED)
   {
          //received a message! Check which one it is, and act
accordingly
```

```
        if (receivedMessage == MES_TAKE_A_BREAK)
        {
              //time to make some coffee. Let's see how far the
sender of this message is away from me,
              //to determine who has to brew the coffee.
              int distance = radio::getDistance();
              if (distance > 100)
              {
                    //too far, I'll brew the coffee myself...
                    hugs::delayGame(3000); //wait 3000 ms
                    //coffee ready, let's inform the nearest
unit
                    radio::sendToNearestDevice(MES_COFEE_READY);
              }
        }
        else if (receivedMessage == MES_COFEE_READY)
        {
              // Drink it.
        }
   }
}
```

OVERVIEW OF ALL RADIO FUNCTIONALITY

Check HUGS_radio.h for detailed information on each function of the radio.

```
void enable();
void disable();
void sendToDevice(HUGs_device theDevice, HUGs_Message mes);
void sendToAll(HUGs_Message mes, const int arg = 0);
void sendToTeam(HUGs_TeamID teamID, HUGs_Message mes, const int arg
= 0);
void sendToNearestDevice(HUGs_Message mes, const int arg = 0);
bool receive(HUGs_Message &theMessage);
int getArgument();
int getDistance();
HUGs_Message registerNewMessage(String messageName);
```

Summary

This thesis proposes a new genre of outdoor games for children, namely Head Up Games. The concept was inspired by the observation that existing pervasive outdoor games for children were mostly played *head down*, as the predominantly screen-based interaction of existing games required constant attention of the children.

First, the vision of Head Up Games is described and illustrated with several design cases (Chapter 2). In contrast to the head down games, Head Up Games aim to encourage and support rich social interaction and physical activity, play behaviors that are similar to play behaviors seen in traditional outdoor games (such as tag and hide-and-seek).

The design process of Head Up Games poses several challenges. In User Centered Design it is commonly accepted to start the development of a new product using low-fi mock-ups, e.g., paper prototypes, and evaluate these with end-users. In the case of Head Up Games this proved to be difficult, as the emerging game experience is significantly altered when using paper prototypes. Therefore, a study was carried out that used high-fi prototypes, i.e. working, interactive, prototypes, from a very early stage in the design process (Chapter 3). This way, the effect of interactions on the game experience can be addressed earlier and better in the design process. Furthemore, having access to technology early in the design process, allows designers to better explore the design space.

However, designers often do not possess adequate skills to quickly prototype interactive products, particularly products that need to be evaluated in an outdoor context. Such a development is often costly and time-consuming.

Therefore, the RaPIDO platform was developed (Chapter 4). The platform not only includes the appropriate hardware for creating outdoor games, but is also bundled with software libraries, to allow designers not specifically trained in software engineering to adopt the platform easily. RaPIDO was evaluated using a case study methodology with two Industrial Design master students. The evaluation not only focused on the usability of the platform, but, more importantly, how the use of the platform affected the design process. The main conclusion of the study was that the designers indeed were able to rapidly create mobile games, and that the hardware offered was suitable for creating outdoor games. Furthermore, issues were identified with regard to writing the game software, e.g., managing the complexity of the software.

Finally, for evaluating Head Up Games with children two methods were applied: the Outdoor Play Observation Scheme (OPOS) was used to quantify the intended play behavior. Furthermore, GroupSorter was developed to provide a framework to interview a group of children simultaneously, resulting in qualitative comments. Both OPOS and GroupSorter were applied for evaluating three Head Up Games, which are described in Chapter 5.

Curriculum Vitae

Iris Soute was born on July 14, 1976 in Stein, the Netherlands. She studied Mechanical Engineering at the Eindhoven University of Technology (TU/e). After graduating in 2001 she took up a full-time position at Philips CFT as a software engineer. She followed a four-month course on embedded, technical software engineering and subsequently worked in projects involving high-precision measurement systems.

After three years of embedded software engineering, Iris decided to make a career switch and she was admitted to the User System Interaction (USI) program at the TU/e in 2004. During the design case project of USI, she and two fellow students developed *Camelot*, an outdoor game for children. *Camelot* provided the inspiration for her PhD research, which she started in 2006 in the User Centered Engineering group in the department of Industrial Design at the TU/e.

List of Publications

JOURNAL ARTICLES

Soute, I., Bakker, S., Magielse, R., and Markopoulos, P. (2013). Evaluating player experience for children's outdoor pervasive games. Entertainment Computing, 4(1), 25–38.

Hooper, C. J., and Soute, I. (2012). TAPT and Contextmapping: Understanding how we understand experience. Literary and Linguistic Computing.

Soute, I., Markopoulos, P., and Magielse, R. (2009). Head Up Games: combining the best of both worlds by merging traditional and digital play. Personal and Ubiquitous Computing, 14(5), 435–444.

PEER-REVIEWED CONFERENCE PROCEEDINGS

Soute, I., Lagerström, S., & Markopoulos, P. (2013). Rapid Prototyping of Outdoor Games for Children in an Iterative Design Process. In Proc. of the 12th Int. Conf. on Interaction Design and Children. New York: ACM. To appear.

Soute, I., Aartsen, H., and Bangaru, C. (2011). On Developing a Platform for Mobile Outdoor Gaming for Children. In D. V. Keyson, M. L. Maher, N. Streitz, A. Cheok, J. C. Augusto, R. Wichert, G. Englebienne, H. Aghajan, and B. J. A. Kröse (Eds.), Ambient Intelligence (Vol. 7040, pp. 200–204). Springer Berlin Heidelberg, Berlin, Heidelberg.

Toering, E., Soute, I., and Markopoulos, P. (2010). Rule customization in head-up games. In Proc. of the 3rd Int. Conf. on Fun and Games (pp. 144–148).

Soute, I., Kaptein, M., and Markopoulos, P. (2009). Evaluating outdoor play for children: virtual vs. tangible game objects in pervasive games. In Proc. of the 8th Int. Conf. on Interaction Design and Children (pp. 250–253). ACM, Como, Italy.

Soute, I., and Markopoulos, P. (2009). Evaluating the effect of the tangibility of game elements on physical activity and social interaction in Head Up Games. In CHI '09 Extended Abstracts on Human Factors in Computing Systems.

Soute, I. (2007). HUGs: head-up games. In Proc. of the 6th Int. Conf. on Interaction Design and Children (pp. 205–208). ACM, Aalborg, Denmark.

Soute, I., and Markopoulos, P. (2007). Head Up Games: The Games of the Future Will Look More Like the Games of the Past. In Human-Computer Interaction – INTERACT 2007 (pp. 404–407).

Verhaegh, J., Soute, I., Kessels, A., and Markopoulos, P. (2006). On the design of *Camelot*, an outdoor game for children. In Proc. of the 2006 Conf. on Interaction Design and Children (pp. 9–16). ACM, Tampere, Finland.

.

INFORMAL CONFERENCE ARTICLES

Soute, I., and Sturm, J. (2011). Towards a Setup for Longitudinal Evaluation of Player Experience with Children. Presented at the 2nd Int. Workshop on Evaluating Player Experience in Games (EPEX11), Bordeaux, France.

Soute, I., Kessels, A., Verhaegh, J., and Markopoulos, P. (2007). Applying “Mission from Mars” for gathering children’s requirements for an outdoor game. In Proc. of the 11th CHI-Nederland Conferentie: HCI Methods, Way to Go? Presented at the HCI Methods: way to go? - 11e CHI-Nederland conferentie.

Soute, I. A. C., Molengraft, M. J. G. van de, and Angelis, G. Z. (2001). Using Genetic Programming to Find Lyapunov Functions. In C. Ryan (Ed.), Graduate Student Workshop (pp. 449–452). San Francisco, California, USA

Dankwoord

A journey of a thousand miles begins with one single step

- Lao Tse

De eerste personen die ik hier wil bedanken, zijn de twee mensen die samen met mij de eerste stap zetten die leidde tot dit proefschrift, al wisten we dat toen nog niet. Angelique en Janneke, samen maakten we *Camelot*, een spel voor kinderen, en *Camelot* werd de basis voor mijn onderzoek. Niet alleen was het een succesvol project, we hebben ook flink wat lol gehad, met marsmannetjes en een roadtrip door Finland inclusief een close encounter met een non-met-Nokia foon... ;) bedankt, dames!

Na die eerste stap volgden er nog veel meer en ik wil hier van de gelegenheid gebruik maken om een paar mensen te bedanken die, kort of lang, met me "mee gereisd" zijn.

Panos, you recognized the potential of *Camelot* and gave me the opportunity to research these types of games further. I learned a great deal from you - up to the point that I hear myself repeating to my students advice that you gave me! Thank you for letting me run my own research - or at least have me believe that I did; I vaguely remember telling you that, "no, I will not create a platform for games". And look what I ended up with... Thank you for sticking by me all this time, I had a great deal of fun, I hope you did too! (and sorry for scaring the hell out of you on one or two occasions ;)

Jean-Bernard, dank je wel dat je eerst mijn eerste en later mijn tweede promotor wilde zijn. Dank voor je frisse en kritische blik op mijn onderzoek; het is de kwaliteit ten goede gekomen.

Thank you, Janet Read, Matt Jones and Emile Aarts for providing valuable feedback on this thesis. Furthermore, many thanks to Narcis Parés and Yvonne de Kort for agreeing to be members of the Doctorate committee.

Herman en Chet, zonder jullie hulp was 'het platform' er nooit geweest, dus een hele grote dankjewel aan jullie! Het was voor mij echt een feestje om met jullie samen te werken en iets moois te creëren - wie weet krijgt RaPIDO in de toekomst nog een vervolg en dan hoop ik dat jullie er ook weer bij zijn!

Dank aan mijn paranimfen, Pap en Saskia. Pap, zo'n 20 jaar geleden had ik de eer om bij jouw promotie je paranimf te zijn en ik nam me toen al voor dat áls ik een keer zou promoveren, ik graag jou aan mijn zijde wilde hebben. Eindelijk is het nu zover! Saskia, dank je wel voor je bijdrage aan dit onderzoek - ik denk hierbij niet alleen aan OPOS, maar ook dat je ervoor zorgde dat ik alle PnP-deadlines haalde ;)!

Hartelijk dank aan alle kinderen, leraren, begeleiders, scholen, en scouting die meegespeeld hebben tijdens dit onderzoek. Dit onderzoek was niet mogelijk geweest zonder jullie medewerking.

Dank aan alle studenten die Head Up Games gemaakt hebben en/of eraan meegewerkt hebben: Saskia, Remco, Eric, Jos, Koen, Guo, Tim, Dirk, Roland en Sussi.

Verder bedank ik al mijn collega's van de UCE groep, en dan in het bijzonder de (ex-)"inwoners" van HG 2.44. Het grote voordeel van in een open-space werken is dat er altijd wel iemand bereid is om mee koffie te gaan drinken.

Complimenten aan Saskia Renders, die ervoor zorgde dat zowel kaft als binnenwerk van dit proefschrift er tiptop uitzien.

Al mijn vriendjes en vriendinnetjes, bedankt dat jullie dat nog steeds zijn. Ik heb jullie de laatste tijd schromelijk verwaarloosd, als het goed is komt daar vanaf nu weer verandering in! Laten we beginnen met een leuk feestje :) Judith, jou wil ik in het bijzonder bedanken - je hebt me op zoveel vlakken bijgestaan: je was impromptu oppas voor Tijl als ik deadlines had; je sleepte me elke week mee naar de sportschool; je zorgde ervoor dat er leuke vacatures in mijn

mailbox verschenen; je sprak me moed in wanneer nodig en je stond op de stoep met een ovenschotel toen Geertjan en ik door de griep geveld waren. Over stoepjes gesproken: zelfs daar draai je, samen met Mieke, je hand niet voor om. Dankjewel voor dit alles (alhoewel dat stoepje inmiddels wel vervangen is ;)

Dank aan mijn familie: mijn ouders, schoonouders, Marcel, Elsbeth, Marjon en Joep. Voor jullie ondersteuning, hulp bij verbouwing en niet in de minste plaats voor alle keren dat ik de kinderen bij jullie kon onderbrengen! Mama en papa, jullie in het bijzonder bedankt voor jullie onvoorwaardelijke steun en liefde. Jullie hebben mij altijd mijn eigen keuzes laten maken en me geholpen als ik dat nodig had. Ik hoop dat ik net zo'n goede ouder word voor mijn kinderen als jullie voor mij geweest zijn en nog steeds zijn.

Marthe en Tijl, zonder jullie was dit onderzoek ongetwijfeld eerder af geweest! Geeft niks; op jullie eigen manier hebben jullie me toch ondersteund tijdens het schrijven van dit proefschrift: Marthe tekende de mooiste kunstwerken en Tijl was zeer vrijgevig met 'nuffeltjes'. Ik hou knetterveel van jullie en kijk er naar uit om weer meer tijd met jullie door te brengen!

In dan is het laatste plekje gereserveerd voor de belangrijkste persoon in mijn leven: Geertjan. Ik zou een hele lofzang kunnen schrijven over jou en wat je voor me betekent, maar ik weet dat je daar niet van houdt (en niet gaat lezen). Ik houd het dan ook bij een spreuk van Loesje, die je me ooit gaf: "Het leven is leuker met jou"- ik zie ernaar uit er volop samen van te genieten. xxis.