Designing a Head-Up Game for Children

Koen Hendrix, Guo Yang, Dirk van de Mortel, Tim Tijs & Panos Markopoulos

User-System Interaction Programme
Department of Industrial Design
Eindhoven University of Technology
Den Dolech 2, 5612 AZ Eindhoven, the Netherlands
+31 40 247 5230

{k.j.j.hendrix, y.guo, h.m.j.v.d.mortel, t.j.w.tijs, p.markopoulos}@tue.nl

ABSTRACT

Head-Up Games [19,20] attempt to combine the technological benefits of modern electronic games with the social and physical advantages of traditional games. To demonstrate this concept, a Head-Up Game for 9- to 11-year-old children was designed and developed iteratively, with intensive involvement of children for play-testing. This paper describes and reflects on the game's design process and the implications regarding the concept of Head-Up Games. The final game, *Stop the Bomb*, was found to be physically and socially stimulating, understood and enjoyed by the target group, and preferred over a non-electronic version of the game at first encounter.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *evaluation/methodology, user-centred design, prototyping*.

General Terms

Design, Human Factors.

Keywords

Head-Up Games, design for children, pervasive technology.

1. BACKGROUND

The video game industry has grown rapidly in the last decade: In the US alone, video game sales multiplied from $ 2.6 billion in 1996 to $ 7.4 billion in 2006 [9]. According to a recent UK survey, 98% of 6- to 15-year-old children play video games, and do this more often than watching movies, calling friends, reading books or magazines, or listening to music [3]. It seems that children prefer electronic games more and more over the traditional, non-electronic, competitors.

This preference for electronic games could well be because modern technology offers toy and game designers new and interesting means to make more appealing toys and games. For example, technology can offer an exciting audiovisual experience, can extend human capabilities (e.g. remote control toys) and can release players of the burden of personally implementing complex game rules (e.g. keeping track of scoring), making it easier and more fair to play [1].

For many this frequent and sustained engagement of children with games is a reason for concern. Traditional outdoor games are being played less often, while they fulfil an important function in the social and physical development of children [17]. In response to such concerns, game manufacturers already offer numerous electronic games affording possibilities for physical and social activity, such as motion-sensitive input devices and a chat channel between online players. However, there still is a considerable gap between the social and physical possibilities of these games and traditional outdoor games. For instance, physical control such as offered by the Nintendo Wii console encourages movement, but does not provide as much physical exercise as a game of tag or the opportunity for rich social interaction that takes place in a game of hide and seek.

1.1 Pervasive games

With the development of pervasive technology, the genre of pervasive games has emerged. Magerkurth et al. define pervasive games as "[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" [11].

Researchers have been developing a number of pervasive games. Examples of outdoor games for adults are *Uncle Roy All Around You* [5] and *Catchbob!* [12]. These games are location-based games, played with a GPS and/or WiFi-enabled mobile devices. An increasing number of pervasive games is introduced not only in related literature, but also as commercial offerings.

To date, the majority of these games is meant for adult players. Fewer examples have been developed of pervasive outdoor games specifically designed for children: two pioneering projects are *Savannah* [4] and *Ambient Wood* [16]. In *Savannah*, children were equipped with PDAs with WiFi and GPS. A virtual savannah was overlaid on a school field, and children had to cooperate as 'lions' to hunt the savannah. In *Ambient Wood*, children took a field trip in a wood that was augmented with mobile and fixed devices that provided contextually-relevant information. Each pair of children carried a PDA and a probe.

The games mentioned above share three characteristics:

- *Location-bound infrastructure.* The game's narrative depends on complex technological devices in an augmented space (e.g. the wood in *Ambient Wood*), or an 'intelligent' playground in other cases. One may expect that in the future both the hardware and the infrastructure will be more readily available, but for the time being such infrastructure is definitely not prevalent, and depending on it for an outdoor game is not consistent with spontaneous play at locations such as schoolyards and parks.

- *Location awareness.* All games assume a form of location awareness. However, such technologies have inevitable imperfections. For example, GPS errors have been shown to adversely affect game play in the evaluations of *Savannah* [4]. Nova et al. [13] report that players that used the automated location awareness tool did not perform better than other participants.
- *PDA as gaming platform.* Interactions with the virtual world are performed through a portable display, that acts as a window to the virtual world.

1.2 Head-Up Games

The characteristics mentioned above may interfere with some of the developmental benefits that are present in traditional outdoor games. To combine the best of both worlds, Markopoulos and Soute [19] proposed the concept of Head-Up Games: a subclass of pervasive games that preserve the benefits of traditional outdoor games while still employing the advantages offered by technology.

The main aspiration for Head-Up Games is that playing them will resemble outdoor games of the past as far as the opportunities for physical activity and social interactions offered to the players, while at the same time bringing the benefits of technology into game play. It is essential that the technology is of added value, to bring some of the benefits of gaming mentioned earlier into the domain of outdoor play.

Markopoulos and Soute proposed some general characteristics to which Head-Up Games should adhere, though they could not report on how well the aspirations for head-up gaming described are met in practice when games are designed accordingly. The characteristics that Head-Up Games should possess are:

- *Simplicity and robustness of technology*, to avoid that its imperfections dominate the gaming experience.
- *Minimal installation*, allowing for play to happen spontaneously in any free space.
- *Head-up focus,* avoiding reliance on visual output on handheld displays, allowing for richer verbal and non-verbal social interaction amongst players.
- *Flexible gameplay*, allowing for players to make their own rules and adapt the game to the social context at hand.

The concept of Head-Up Games shares some motivations with, but is not identical to, the notions of Exertion interfaces [12] (which can be confined to indoor play), intelligent playgrounds [18] (which concern a specific location bound infrastructure), or social games [7][14] (which largely ignore the motivation for outdoor physical activity).

The remainder of this paper discusses a game for 9- to 11-year-old children that was developed adhering to these principles. The design process, the game concept, its implementation and its evaluation are presented. The paper concludes by reflecting on the extent to which the aspirations for Head-Up Games have been achieved with this design and lessons learnt from this first attempt to design a Head-Up Game (in the sense that it was defined above).

2. GAME DESIGN

A user-centred iterative game development process was adopted, which is illustrated in figure 1. First, a number of initial game concepts were developed and prototyped. Based on a playtest of these prototypes, some games were abandoned and others improved. The remaining games were playtested with

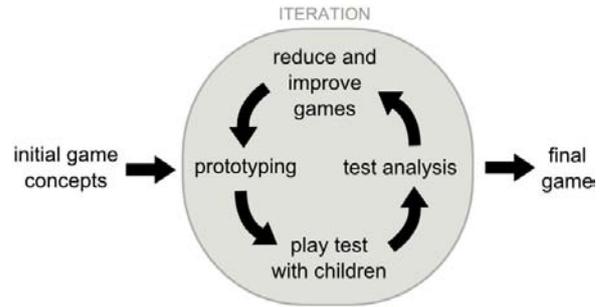children again. In this way, the initial game concepts were gradually narrowed down to a single game and prototype.

Figure 1: Game development process.

In section 3 of this paper, the first iteration is described: creation of four game concepts and the creation and evaluation of paper prototypes. Section 4 describes the second iteration, in which the number of games was reduced to two, and limited-functionality prototypes were created and evaluated. In the third iteration a more advanced prototype of one final game was created, and this 'high-tech' prototype was compared to a paper-based variant as described in section 5. Finally, in section 6, we reflect upon the applied game design methods, and discuss the Head-Up Games concept in general.

2.1 Initial game concepts

To bootstrap the design process, two semi-structured interviews were held with two primary school teachers. The interviews were focused on current playground activities of 9- to 11-year-old children, and potential issues of play testing with children. In a series of brainstorm sessions, the authors generated ideas around the theme of Head-Up Games. The focus of these sessions narrowed down from children's play activities in general to specific Head-Up Game concepts. The brainstorms resulted in nine initial game concepts, which were documented and critically reviewed by the authors in terms of feasibility, adherence to the Head-Up Game characteristics, and suitability for the target group. In the end, 4 game concepts were retained, for which low-tech prototypes were created for play-testing with children.

Figure 2. Four games being played in the first test. From left to right: Africans & Lions, Timeball, Pet Care and Defuse the Bomb.

Africans & Lions is an individual game where each player is either an African or a Lion; this role is not visible to the player himself/herself, but is visible to all other players. Because everyone knows the number of Lions and Africans in the game, a player must find out his/her role by checking the role of the others on other players' backs. Lions can then survive by capturing an African, while Africans need to get water from a well (a designated area) while avoiding Lions to survive. The surviving players are the winners of that round.

Timeball is played by two teams each having its own goal. There are two balls in play at the same time, which can be used by any team. Both teams have to score by throwing an activated ball into the opponent's goal. A ball can be activated at one's own goal, and will remain activated for 30 seconds. Players have to pass the ball to team members and are not allowed to run or walk while holding a ball.

Pet Care is played in two teams that each have a pet animal to feed. Each animal wants a random list of food items in a particular order. Players have to gather food items, which are scattered across the playground, and bring them to their pet to find out whether it is the animal's desired food at that time. The first team to feed its animal its entire list of food items wins. Players can only carry one food item at a time, but trading food with the other team is allowed.

Defuse the bomb is played in two teams: Mafia and Police. The Mafia have activated a bomb that will explode in three minutes. To win the game, the Police must defuse the bomb by quickly obtaining four keys that are initially in the Mafia's possession. Police can arrest Mafia players, and take their keys, by tagging them. Mafia players try to avoid being caught and are allowed to throw or hide keys. If the bomb explodes, the Mafia team wins.

3. INITIAL CONCEPT EVALUATION

To evaluate whether our four initial game concepts were fun and suitable for children, as well as to measure their adherence to the Head-Up Games characteristics, paper prototypes were created and play-tested with 13 children.

3.1 Low-tech prototypes

For *Africans & Lions*, pictures of the player roles (Africans and Lions) were printed on cardboard and attached to a cord. Before starting the game, players were instructed to close their eyes, allowing referees to hang a card on each player's back. Different game areas were marked on the field using coloured cones.

The *Timeball* prototype consisted of football-sized foam balls. The activation time and score were kept track of by the test administrators (referees), and a de-activation of the ball was announced to the players in the field. Plastic laundry baskets were used as goals.

The food items in *Pet Care* were made of printed cardboard cards and were spread across the play court in small baskets. The pets were represented by two cuddly toys ca. 50 cm in height. Behind each cuddly toy sat one referee who acted as the voice of the pet.

For *Defuse the bomb*, the bomb was made from a decorated cardboard box with four holes for the keys, and an inexpensive kitchen timer. Four rubber balls (tennis ball size) were used to represent the bomb's keys.

3.2 Expectations

All games were new to the children, which we expected would lead to some confusion and ineffective actions which would, in turn, probably limit the children's enjoyment. It was expected that the children would have the most problems with understanding Africans & Lions, because having a hidden status on one's back is a rather unfamiliar game element.

It was also expected that children would show a lot of physical activity, such as running around and throwing balls, because the games' rules strongly encourage this. It was expected to see at least basic social interaction (such as shouting to other players), but because the games were new to the players and only played shortly, we did not expected to see many instances of richer social interactions (e.g. trading or discussions of team tactics).

3.3 Test method

Two methods were used to evaluate to what extent the games conformed to the Head-Up Game principles and whether the game was suitable for the target group. First, every test was videotaped with one camera positioned at a corner of the playing area, and the video was qualitatively analyzed afterwards. Second, the children were asked to complete a questionnaire after each game they played.

Coding was done by two observers in the first two tests and four observers in the final test, using time sampling and observing all children at play (rather than focusing on one child). The four aspects evaluated in the video analysis (adapted from [2]) are complexity, fun, social interaction, and physical activity. In table 1, the different observed behaviours corresponding to these four criteria are summarized. Complexity was rated by looking for wrong actions and for expressions of doubt. Fun (or lack thereof) was evaluated by looking for expressions of joy or passivity. Various forms of communication between players, not necessarily including speech, were deemed expressions of social interaction. Physical activity was rated based on the amount of time children engaged in active physical movement; merely walking was not included. Naturally, wrong actions, doubt and passivity should be as low as possible; joy, social interaction and physical activity should be high.

Table 1. Video evaluation scheme.

Criterion	Behaviour
Complexity	Wrong actions: **Ineffective actions or actions not allowed by the rules**
	Doubt: **Looking around for help or asking questions**
Fun	Passivity: **Withdrawing from the game, not playing anymore**
	Joy: **Engaged play, laughter or winning gestures**
Social interaction	Social interaction: **Talking, shouting, passing objects, or trading objects**
Physical activity	Physical activity: **Running, throwing, climbing or kicking**

In every video analysis, each observer rated each of the six behaviours for every round of play. A rating was given on a 4 level scale (none–low–medium–high). All ratings for a game were then compared and averaged. The final rating for a game

was determined by reaching consensus between all observers. This analysis procedure was deliberately coarse-grained to help select between the prototypes and capture large differences in the observed play patterns.

Besides the video analysis, children's appreciation of the games was also measured by surveys. After playing a game each child was given a short questionnaire consisting of a list of eight to ten aspects of the game, which included items like 'the bomb', 'activating the ball', 'running away', 'being Police', 'working together' and 'scoring a goal'. Each game had a different list of aspects relevant to that game. The children were instructed to mark both the best and the worst aspect of the game. In addition to this measurement, the children discussed their experience of the game in small groups (3 to 4 persons) with one test administrator. After the last game had been evaluated, all children and the experimenters joined a final group discussion to reflect on the children's experiences with the games and discuss possible improvements.

As an experiment in participatory game design with children [8], children were asked for game design input during the playtest. Each round we asked players to collaboratively decide on changing one rule of the game after the second round of play.

3.4 Participants and procedure
To prepare for the first end-user test, a pilot test involving 7 adult participants (colleagues) at the Eindhoven University was held. This pilot provided a number of practical insights regarding the test procedure and some first insights into players' experience of the games. Apart from some minor adaptations to the test procedure, no conclusions were drawn from this preliminary try-out.

For the first end-user test, a test session with children was organized at a child day-care centre in the Netherlands. A total of 13 children aged 6-12 play-tested all four games outdoors in cold dry weather. A number of children quit the test after the first game, possibly because of the cold weather, so later game sessions were played with fewer children. As not enough children of the right age were present, some more children were invited to join who were slightly younger (6-9 years).

The test was held outdoors at the day-care's playground (ca. 30x20 meters). The test lasted around 2 hours in total, in which each game got ca. 10 minutes of actual play time, in which 3 rounds were played.

3.5 Results
Overall, even though the participant group changed in between games, the games were enjoyable for the children had interesting ideas how to improve them. The youngest children (who were younger than we were prepared for) were unable to individually fill in the questionnaires, while others simply copied each others work or encircled all options. Therefore, it was decided not to use the data from the best/worst aspects questionnaire. The video analysis (see table 2) showed that *Timeball* worked very well, having both a lot of social interaction (in the form of passing) and physical activity.

For *Defuse the Bomb*, the video analysis revealed that hiding the keys and searching for hidden keys prohibited social interaction in this game to some extent. The fact that there were as many keys as Mafia players lead to individual play (little teamwork). Besides that, the rule of handing over keys occasionally caused some trouble: A number of times, a Mafia

player did not feel like handing over their key after being arrested (tagged), which led to rough play.

Pet Care was less enjoyable to the children, the continuous running tired them out quite quickly, and they did not engage in trading food items at all.

Regarding *Africans & Lions*, there was some doubt whether the game would be too complex for children. The playtest showed that players who belonged to the target group managed to understand this aspect of the game sufficiently, but only after playing it a few times, at which point they actually liked this aspect. In general, we observed that – contrary to our expectations – the children picked up on the games quite quickly.

Table 2. First test video evaluation results.

	Africans & Lions	Timeball	Pet Care	Defuse the bomb
Wrong actions	H	L	L	M
Doubt	M	L	L	L
Passivity	L	L	M	L
Joy	H	H	M	H
Social interaction	H	H	M	M
Physical activity	H	H	H	H

positive results — negative results

In final group discussion, most children also indicated they least liked *Pet Care*. They experienced the gameplay as repetitive, and did not enjoy continually running back and forth to obtain a food item without knowing whether the animal would want that food. There was no agreement among the kids about which of the other games was the best. The children also reported that they understood the games, although some still had some questions about *Africans and Lions*.

Inviting children to change a game rule between rounds was less successful than we had hoped. The participants had many ideas for improvement, but these were aimed more at improving their recent performance in the game rather than improving the game as a whole. For example, even when the police team had won *Defuse the Bomb* every time, police players still proposed rules to make the police's task easier. Furthermore, discussing and deciding upon a single rule change in between rounds took too much time.

4. INTRODUCING TECHNOLOGY IN THE GAME
After the evaluation of the low-tech prototypes, the number of games retained in the second iteration was reduced down to two. *Pet Care* was abandoned completely since it was insufficiently appreciated. *Timeball* was kept unmodified since it appeared largely appealing without major weaknesses. *Defuse the Bomb* and *Africans & Lions* were merged into one game called '*Stop The Bomb',* aiming to combine the best

aspects of both games: the discovery part of the former and the cooperative aspects of the latter. In *Stop the Bomb*, players are secretly and randomly assigned to Police or Mafia. After finding out who they are, they can start retrieving or hiding the key (as in *Defuse the Bomb*). There is just one bomb key in *Stop the Bomb*, as multiple keys led to individual play.

4.1 Prototypes

In this iteration, some features of the two game prototypes were implemented technologically, while others were simulated using a Wizard of Oz approach.

For *Timeball,* two foam balls were enhanced with red and green LEDs, connected to the re-purposed innards of radio-controlled toy cars. Referees could activate the lights using the toy's original remote control. The prototype is shown in figure 3. Two wicker baskets (placed on the ground) were used as goals.
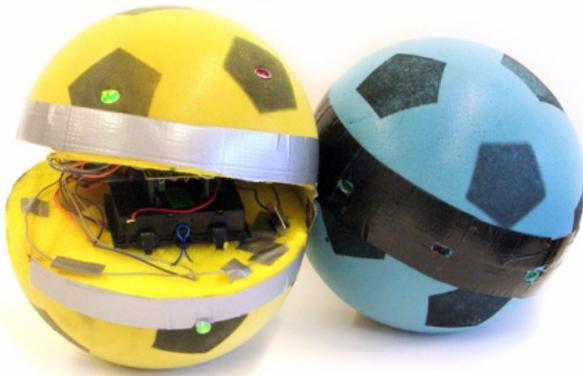

Figure 3. Timeball prototype.

For *Stop the Bomb,* the assignment of roles was digitally implemented in the prototype. Ten belts were constructed, each with a back pocket for a Crossbow Mica2 wireless mote[1] and six LEDs (3 green, 3 red), as shown in figure 4. The belts themselves were made with strong seat-belt-like cloth, and could be sized to fit with Velcro. The bomb key was still a simple rubber ball.

The wireless motes form a network and communicate with a Programming Board, which is connected to a laptop and facilitates control of the LEDs on each individual belt. This allowed for remotely assigning random roles to the players upon game start.

A small number of adaptations had to be made the rules from predecessor *Defuse the Bomb* to make the game work. As players did not know their team at game start, the key was not given to the Mafia players anymore, but was instead laid down at a central place on the playground. Players of both parties were now allowed to pick up the key first, and both parties were allowed to continuously intercept the key from the other by tagging. As soon as a player discovered his/her role, he/she had to scream aloud that role (e.g. "I am police!") before being allowed to pick up the key. In case a player was wrong (e.g. a policeman saying "I am mafia!"), that player was taken out of the game for that round.

The bomb was represented by an electronically-enhanced cardboard box with a slot for the ball. When the key (ball) was put into this opening on time, a Phidget[2] switch was activated,

which triggered an applauding crowd sound sample from a laptop and speakers hidden inside the box. In case the police did not manage to bring back the key in time, an explosion sound was triggered on timeout, indicating the Police's failure and thereby Mafia's victory.
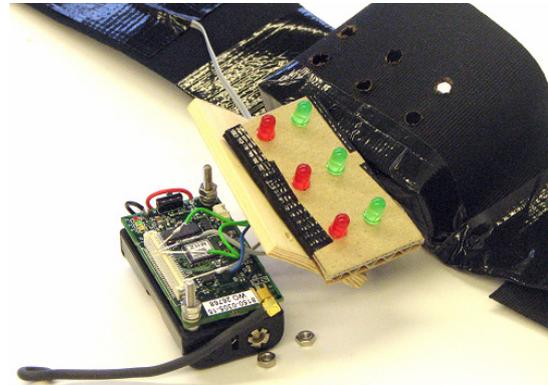

Figure 4. The belt for *Stop the Bomb*, with the back pocket disassembled to show parts (above) and worn by children during play (below).

4.2 Expectations

For *Stop The Bomb*, the combination of role discovery with tagging in teams should make the game more appealing since these two aspects were appreciated most in the preceding game ideas. The fact that *Stop The Bomb* involves both role discovery and tagging was expected to generate richer social interaction (through team communication, hiding, observing others) and physical activity. In addition, the technology (LEDs) should increase the players' joy compared to the low fidelity prototype.

For *Timeball*, it was expected that the installation of remotely controlled LEDs on the ball would make the game more attractive, as well as make the ball's status (not activated/ green/red) more apparent and thereby smoothen the gameplay.

4.3 Evaluation method

The two games were evaluated using the same method as for the low-tech prototype, discussed above in section 3.3. Although the questionnaire method did not work well in the first iteration, it was retained because the participants of this iteration would be of the right age group (whereas some previous participants were too young).

---

[1] Crossbow Technology Inc., www.xbow.com

[2] Phidgets Inc., www.phidgets.com

In the video analysis, we now also allowed intermediate ratings: 'low-medium' or 'medium-high'. We did not ask children for a rule change anymore. In the final group discussion, all children were asked to give a forced-choice preference for one of the two games. Two observers scored the video.

4.4 Participants and procedure

In view of the age group problems at the day care centre during the first test, a second end-user test was held at a primary school in The Netherlands. One class of 21 children (12 girls, 9 boys) aged 10-11 participated. Participants were split into 3 groups of 7 children. *Stop the Bomb* was played once by each group. While waiting for their turn, the other groups sat down on benches along the court, watching the others play. For *Timeball*, two teams played against each other while the third team rested.

Unfortunately, a few of the *Stop the Bomb* belts had broken down during transport, and had to be replaced by cardboard cards before the game started. The combination of using both belts with LEDs and belts with cards may have adversely affected the children's experience of the game.

4.5 Results

The video evaluation results are shown in table 3. Both games created sufficient physical and social activity. *Timeball* worked very well, producing a large group of engaged, shouting children. The role discovery part of *Stop the Bomb* was still a bit confusing, which might have been caused by the inclusion of cardboard cards, but it was appreciated and sufficiently understood after a few rounds. Children seemed to enjoy the colour changes of the ball in *Timeball*, as well as the LEDs on the belt and the audio effects in *Stop the Bomb*.

Interestingly, although the video evaluation indicated more problems with *Stop the Bomb*, the forced-choice game preference in the group discussion revealed that 18 of the 21 participants preferred it over *Timeball*.

Table 3. Second test video evaluation results.

	Timeball	Stop the Bomb
Wrong actions	L	L-M
Doubt	L	L-M
Passivity	L	L
Joy	H	M-H
Social interaction	H	M-H
Physical activity	M-H	M-H

positive results → negative results

The best-rated aspects of *Timeball* were 'activating the ball' and 'balls with lights' with 5 votes each, the worst aspect was 'time pressure' with 6 votes. The best aspect of *Stop the Bomb* was 'discovering what you are' with 5 votes. Interestingly, the

worst features also included 'discovering what you are', along with 'working together' and 'being Police' with 3 votes each. A few children only marked a best aspect.

However, a new major flaw in *Stop the Bomb* was revealed. Because both teams were now allowed to tag each other to intercept the key, back-tagging disturbed the flow of the game. The game regularly got into a stalemate with many players of both teams standing close together and continuously back-tagging each other, leading to quarrels about a back-tagging rule.

Additionally, the secret identity on players' backs inhibited team play to some extent. We expected that players who discovered their identity would inform their team-mates of their status, but instead the first discoverer would simply grab the key and run off. While some children would run after the ball, others were still trying to find out to what team they actually belonged. Sometimes the game would end with some children still not knowing whether they were on the winning team or not.

Although table 3 shows that Timeball scored slightly better on the video evaluation criteria, both games were considered to be good Head-Up Games. Because *Stop the Bomb* was preferred by the children as well as considered to be more innovative by the design team, it was decided to continue working with the game *Stop the Bomb* for the next and final iteration.

5. EVALUATING THE ADDED VALUE OF TECHNOLOGY

Two different versions of *Stop the Bomb* were compared in the third iteration: A 'paper' version using cardboard status cards and a ball, and a 'technological' version that uses LEDs and a key that is represented digitally using vibromotors. The test served to investigate whether and how a digital key could replace a ball, and whether this would make the game more or less suitable for our design goals.

5.1 Prototype

The main adaptation in this iteration was the digital implementation of the key. By sending messages between all wireless motes and using the motes' built-in signal strength indicator, a proximity-sensing mechanism was programmed into the motes. This proximity-sensing has limited precision because signal strength depends on more factors than just the distance between the motes themselves (such as antenna orientation, physical obstacles, and humidity). A small vibromotor was fastened on the inside of each belt so that the wearer could feel the vibration on the front of his/her hip. The vibromotor was connected to the belt's wireless mote; vibration in the belt was used to indicate possession of the key.

On game start, roles are randomly distributed. Five out of ten belts receive the Police status (green LEDs light up) while the other five belts become Mafia (red LEDs). Then, in one of the Mafia players' belts, the vibromotor starts vibrating, indicating to the player that he/she possesses the key. That player will also immediately know he or she is Mafia since the key is always initially given to a Mafia player. As soon as another player gets close enough to the key-carrier, the vibration stops in the first belt and turns on in the newcomer's belt. Because the proximity-sensing mechanism is rather crude, the distance at which the key will 'jump' can differ from a few feet to a few yards. This randomness was accepted as a part of the game. To prevent the digital key from rapidly jumping back and forth between two adjacent belts, a player will always keep the key for at least three seconds, no matter how close other belts are.

Because the key is now a digital signal, the bomb should also be able to receive that digital signal and defuse itself if someone comes close enough with the key. However, due to time constraints this was not implemented. For the technological bomb we used the previous iteration's cardboard box again. Since there was no physical key, the Phidgets mechanism and audio effects could not be used.

Figure 5. Two versions of Stop the Bomb compared in the third test: a 'paper' version using cardboard cards (above), and a technological version using belts with a wireless mote, LEDs and a vibromotor (below).

5.2 Expectations

The implementation of the digital key was expected to change the gameplay in two major ways. First, role identification will be faster since one of the players (the key-carrier) will immediately know his role. It is expected that this player will directly gather as many Mafia team-mates as possible by informing them of their status, and then discuss a team strategy together. Second, the digital key-passing mechanism will be rather random due to technological limitations of the proximity sensing method. This randomness might make the game more appealing to weaker players, but less fun for stronger tactical players.

5.3 Evaluation method

The two games were evaluated using the method from the earlier sessions. Again, we allowed intermediate ratings 'low-medium' or 'medium-high' in the video analysis. The video was scored by four observers.

Apart from the best/worst aspect questionnaire, all participating children were given an additional form after they had played both versions of the game. Children were asked to rate their preference for either version of the game, on a five-point bipolar scale from "strongly prefer paper version" to "strongly prefer technological version". They were also asked for their opinion on the randomness of the digital key, on a five-point scale from "Very stupid / unfair" to "Very nice / surprising".

5.4 Participants and procedure

The third end-user test was held in a sports hall (ca. 20x25 m.) at a primary school in the Netherlands (different from the previous test's school). One class of 23 children (8 boys and 15 girls, aged 9 to 11) participated. The sports hall was split into two fields, and the class was divided over these two fields. After an introduction and explanation, one group played the paper version while the other group played the technological version. After 10 minutes of play, the game was evaluated and the groups switched fields. This allowed every child to play both versions of the game. The entire test lasted around 50 minutes.

5.5 Results

The results of our video analysis are shown in table 4. The technological version was more complex than the paper version, and it led to many questions during the first round, but the game became less confusing and more enjoyable to the kids in subsequent rounds of play.

Table 4. Third test video evaluation results.

	Paper	Technology
Wrong actions	L	L-M
Doubt	L-M	M
Passivity	L-M	L
Joy	M-H	H
Social interaction	M-H	M-H
Physical activity	M-H	M-H

positive results — negative results
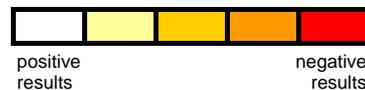
In terms of social interaction and physical activity, video analysis showed there was little difference between the paper and technological version. Of course, there was no key throwing in the technological version as there was no physical key, but the absence of throwing led to more running and chasing because since both team-mates and opponents now had to close in on each other to receive the key. While sometimes all players would stand still to find out who had the invisible key, everyone would start running as soon as the key-carrier was discovered. The following quote from children playing the technological version illustrates this type of game play:

[A group of six children is running after child B]

A: "Run! Get her!"
B: "I lost it!"

[B slows down and the group comes to a standstill, everyone is looking at each other suspiciously]

C: "Someone has to have it."
D: "Who has it now?"
E: "You have to feel it, here!"
D: "I don't have it!"
F: "Me neither!"
A: "Who has the key!?"

[Player D suddenly runs off]

D: "I have it!!"

[The group starts chasing D]

Our discussions with the children confirmed that the vibrating belts were a big hit with the children. The children were anxious to try the belts; they found them exciting and did not mind wearing them. To quote one child at the end of playing the technological version:

Child: "Are we finished?"

Referee: "Yes."

Child: "Oh, too bad! Oh, this was such a fun game." *[to newcomers:]* "Hey guys, this is so fun, this is so cool!"

In general, children liked both versions of the game a lot. According to the questionnaires, the best feature in the technological version was "the key", again showing how much the children liked the vibrating belts. In the paper version it was "discovering your role". In both versions "the bomb" was voted as the worst feature; possibly because the bomb was just represented by an empty cardboard box. Adding sound effects to the bomb, which was appreciated in the previous iteration, might make the technological version more exciting.

The responses to the final two questions, which were two five-point-scale ratings, are tallied in table 5. 17 out of 18 children indicated they strongly preferred the technological version over the paper one (mean 4.8, SD 0.9), and 16 out of 18 thought the randomness in the invisible key was a positive feature (mean 4.0, SD 1.0). Three children did not fill out the form.

Table 5. Comparison between paper and electronic version. Tallied five-point scale ratings.

Which game did you like best?				
Strongly prefer paper version	.	.	.	Strongly prefer technological version
1	0	0	0	17

What did you think of the randomness of the key?				
Very stupid / unfair	.	.	.	Very nice / surprising
1	0	1	1	15

6. DISCUSSION

Below we reflect on lessons learned about the design approach used and about the general theme of Head-Up Games.

6.1 Head-Up Games concept

This design study has demonstrated the potential of Head-Up Games. Children enjoyed the final game, and the electronic enhancements had a demonstrable added value. The game created also seemed to establish roughly as much social interaction and physical activity on the playground as an average traditional outdoor game, but this statement is based only on very coarse observations and comparisons. More refined structured observations and a formal experiment are needed to assert with confidence that Head-Up Games evoke social interaction and play patterns are equivalent to traditional playground games. Play behaviours of Head-Up Games should also be compared to pervasive games relying on handheld displays.

The adopted iterative design process exposed the characteristics of Head-Up Games described above to the children as elements of the prototypes. The aspiration to emulate traditional playground play seems valid and feasible. Furthermore, the evaluations with the children suggest that the avoidance of visual displays, as advocated in [19], needs to be relaxed. We would like to argue that the focus in designing Head-Up Games should not be on avoiding visual displays at all costs; visual displays have a great potential for delivering exciting play experiences. Occasionally glancing at a screen does not automatically destroy a game's allowance for rich social interaction. Head-Up Games should be fun games that are movable, scalable and allow for physical activity and spontaneous social interaction. Visual displays exist in many forms, and more adventurous solutions than LEDs or handheld displays ought to be explored, as they might be very suitable for Head-Up Games.

6.2 Asking children for improvements

It was reported above how our attempt at participatory design [8] by asking children for a rule change during play did not turn out to be a good source for design ideas and improvements. However, it remains an interesting indicator of how children experienced the game.

Eliciting new ideas for improvement by means of a group discussion after all games appeared to be more valuable. Once a number of suggestions have been made in the classroom, many kids think of new alterations. This led to a long-winded but very valuable group conversation where many game issues and improvements were discussed.

6.3 Evaluation methodology

Regarding the involvement of children in the design process, using survey techniques as advocated in related literature (see for example [15]) was found ineffective. Children were concerned with how they could have won the game rather than evaluating how much fun the game was for the whole group and how to improve the game as a whole. This might be a recency effect [10]; a child who just lost two successive rounds is more likely to blame his/her team and consider "teamwork" to be the worst aspect, and a child who has just won twice while being Mafia might consider "being Police" the worst aspect. In this respect, surveying children about a game where they can win or lose differs from surveying them about an input technique or an educational application. This consideration must be taken into account when preparing questionnaires or interviews for evaluating game play.

The video analysis scheme we employed is focused on four main aspects of the game: Complexity, fun, social interaction, and physical activity. Marks ranging from 'high', 'medium', 'low' and 'none' were used to evaluate these aspects of the game mentioned above. However, this scheme was too coarse-grained to make subtle distinctions between different types of social interaction. Interactions that would be considered 'medium' judging by their intensity or duration might differ very much in terms of quality and in terms of the social skills they help children practice (and hopefully develop). Also, there is no clearly observable threshold for helping determine between head-up and head-down behaviours in treating game objects; e.g., would it be head-down if you glance at a screen or if you focus on a game object like a ball? Similarly, how should the coder decide what exact amount of physical activity should be considered 'high' or 'low' for Head-Up Games?

In order to support structured observation, it is necessary to develop a detailed coding scheme for play activities that includes multiple, well-defined aspects of social interaction and physical activity. Possibly, other measures than video analysis (e.g. physiological measures to measure activity [6]) might be useful to evaluate and compare games as well.

7. CONCLUSION

We succeeded in designing a game that adheres to the Head-Up Game principles. Using video analysis and questionnaires, we have shown that the game *Stop the Bomb* is fun and not too complex for 9- to 11-year old children, and that the game promotes both physical activity and spontaneous social interaction. It was also shown that, when playing the game for the first time, children preferred the technological implementation of the game to a non-technological alternative. This indicates the potential of integrating pervasive technology in outdoor play.

The design case presented demonstrates the feasibility of the conception of Head-Up Games and the potential of this new type of games. However, this work has been limited to first-experience tests, and studies of long term usage are required as further evidence of the credibility of the concept.

8. ACKNOWLEDGEMENTS

9. REFERENCES

[1] Adams, E. & Rollings, A. (2007). *Game Design and Development; Fundamentals of Game Design.* Pearson Education, NJ.

[2] Barendregt, W. (2005). *Evaluating Fun and Usability in Computer Games with Children.* PhD Thesis, Eindhoven Universtity of Technology, the Netherlands.

[3] BBC New Media Research (2005). *Gamers in the UK: Digital play, digital lifestyles.* Available at http://open.bbc.co.uk/newmediaresearch/.

[4] Benford, S., Rowland, D., Flintham, M., Drozd et al. (2005). *Life on the edge: supporting collaboration in location-based experiences.* CHI '05, pp. 721–730. ACM Press, New York.

[5] Benford, S., Seager, W., Flintham, M., Anastasi, R., et al. (2004). *The Error of Our Ways: The Experience of Self-Reported Position in a Location-Based Game.* In:Mynatt, E., Siio, I. (eds.) UbiComp 2004. LNCS, vol. 3205, pp. 70–87. Springer, Heidelberg.

[6] Boogaard van den, C.M.H. et al (2007). *Bewegen met computergames.* KvL/B&G 2007, 113, TNO, the Netherlands.

[7] Brederode, B., Markopoulos, P., Gielen, M., Vermeeren, A., and de Ridder, H. (2005). *pOwerball: the design of a novel mixed-reality game for children with mixed abilities.* Proceedings of IDC '05 (Boulder, Colorado, June 08 - 10, 2005). ACM, New York, NY, 32-39.

[8] Druin, A. (2002). *The role of children in the design of new technology.* Behaviour and Information Technology, 21 (1), 1-35.

[9] Entertainment Software Association (2007). *Essential facts about the computer and video game industry, 2007.* http://www.theesa.com/facts/sales_genre_data.php

[10] Murdock, B. (1962). *The serial position effect of free recall.* Journal of Experimental Psychology, 64, p.482-488.

[11] Magerkurth, C., Cheok, A.D., Mandryk, R.L., Nilsen, T. (2005). *Pervasive games: bringing computer entertainment back to the real world.* Computer Entertainment, 3 (3), p.4, ACM Press, New York.

[12] Mueller, F., Agamanolis, S., and Picard, R. (2003). *Exertion interfaces: sports over a distance for social bonding and fun.* In Proceedings of CHI 2003 (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). ACM, New York, NY, 561-568.

[13] Nova, N., Girardin, F., Molinari, G., Dillenbourg, P. (2006). *The Underwhelming Effects of Automatic Location-Awareness on Collaboration in a Pervasive Game.* International Conference on Cooperative Systems Design, pp. 224–238.

[14] Piper, A.M., O'Brien, E., Morris, M.R. and Winograd, T. (2006). *SIDES: a cooperative tabletop computer game for social skills development.* Proceedings of the CSCW'06, (November 04-08, 2006, Banff, Alberta, Canada). ACM, New York, NY, 1-10.

[15] Read, J. C. and MacFarlane, S. (2006). *Using the fun toolkit and other survey methods to gather opinions in child computer interaction.* Proceedings of Interaction Design and Children 2006 (Tampere, Finland, June 07 - 09, 2006). IDC '06. ACM, New York, NY, 81-88.

[16] Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., et al. (2004). *Ambient wood: designing new forms of digital augmentation for learning outdoors.* Proceedings of IDC '04, pp. 3–10. ACM Press, New York.

[17] Scarlett, W.G., Naudeau, S., Salonius-Pasternak, D. & Ponte, I. (2005). *Children's Play.* Sage Publications, CA.

[18] 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 (Montréal, Canada, April 22 - 27, 2006). ACM, New York, NY, 303-308.

[19] Soute, I. & Markopoulos, P. (2007). *Head-Up Games: The Games of the Future Will Look More Like the Games of the Past.* Lecture notes in Computer Science, vol. 4663, p.404-407. Springer.

[20] Soute, I. (2007). *HUGs: Head-Up Games.* Proceedings of IDC'07, p.205-208. ACM.