ABSTRACT

Pervasive and location-based games are played in the real world rather than on the screen of a computer or mobile device. This makes them difficult to study. Since players move around it is difficult to observe them, while at the same time many of the central game activities cannot be monitored simply through logs of device interaction. In our project, we develop tools that allow players to record their subjective experiences during an ongoing game. We report on the design considerations for such tools, and our first experiences of using them in a game session.

Categories and Subject Descriptors

H.5.2 User Interfaces: Evaluation/Methodology
I.1.4 Computing Methodologies: Applications

General Terms
Documentation, Design, Experimentation

Keywords
Pervasive games, self-reporting, evaluation, emotion, experience, activity

1. INTRODUCTION

Games played in the physical world are not videogames. They easily turn into pervasive games - lived experiences where game and life merge, and where coincidences such as fighting weather conditions, being stopped by a policeman, or picking up a game-related flyer often are reported as the key moments of the game. They are games where the individual experiences of players are shaped by physical activities and face-to-face meetings with people in and out of the game. Understanding and documenting the experience of such games is far from simple. Controlled lab studies and post-game questionnaires take us very far from the lived experience, which means that we need new methods to study pervasive games.

In this paper, we report on the design of a hand-held self-reporting tool, intended to capture aspects of the game experience during a pervasive game event. The experience captured is then discussed with the players in a post-game debrief. We describe the design rationale behind the tools, the design of the tools and the post-game debrief system, and our experiences of using the tools to evaluate a pervasive game.

2. BACKGROUND

2.1 Understanding play experience

Although computer games have been subject of study since the eighties, studying the game experience as such is a relatively new phenomenon; an aspect of a recent trend in human-computer interaction studies to move away from traditional usability towards experience evaluation [18]. Being intrinsically motivated, games can be seen as a form of aesthetic experience [8]. Digital games need to be usable, but at the same time they should not be easy to play: the risk of failing is an intrinsic part of playing. Csikszentmihalyi [7] identified the concept of flow as central to play: the ultimate balance of challenge and competence into an absorbing experience.

The fact that games are centred on activity and absorption creates particular challenges for studying them. It is often undesirable to disturb the players while gaming, to ask them about what they experienced or why they chose to act in a certain way, as this breaks the game experience. At the same time, interviews and surveys that are carried out after a game session are heavily influenced by how it ended. Experiences of confusion, frustration and stress can be redeemed by the fact that the players eventually succeed, and de-emphasised or left out from post-game interviews. The problem is even larger for games that have a role-play or competitive aspect, where players are not only expected to play but also to perform. In such games, players may report the game as being easier or more engaging than what they actually experienced, as they otherwise might appear as less skilled players. This which we may call the post-game lie, can severely hamper the reliability of subjective experience reports gathered in post-game interviews and surveys.

2.2 Studying pervasive games

Pervasive games are games that extend beyond a given time and place and set of players to merge with ordinary life [20]. They are typically played in vastly distributed areas in the physical world, and they can sometimes extend over long periods of time. Pervasive games are typically not web services or applications running on a computer or mobile phone, but technology-supported activities in the real world that rely on a mix of devices, programs and special-built technology.

This poses major challenges for capturing the player experience. If the game technology supports some but not all of the game activities, technology logs are not sufficient to capture what players actually do. If players move over large areas it...
becomes almost impossible to capture their activities on video. The long duration of game sessions tend to emphasise the post-game lie, as it becomes hard for players to even remember what they felt and did early on in the game.

Previous of pervasive games have largely focussed on post-game interviews and surveys completed by system logs [15]. Some game studies have tapped into the in-game communication between players [2,10] and between players and game masters [25,26]. The latter gives better insights into the actual game experience than activity logs do, but it is only possible in games where communication can be monitored through the system. Installing fixed or mobile surveillance equipment into the game [25] is more rarely used and faces ethical, legal and technical difficulties – while still failing to capture the subjective experience of players. Jonsson et al. [14] report in-game diaries as a valuable resource in detecting periods of confusion, boredom or stress. However, this alternative is only feasible for slow-paced games with extremely long duration.

3. METHODOLOGICAL CONTEXT

The goal of our project is to develop self-reporting tools that can capture something of the players’ game experience during an ongoing game session, primarily to support post-game interviews. In this endeavour, the perhaps most central design decision is to define what parts of the gameplay experience that the tool should capture. We decided to develop two tools, directed towards capturing different aspects of the in-game experience.

3.1 Models of Immersion and Flow

If possible, measuring ‘flow’ would be desirable. Csikszentmihalyi’s original concept is however rather elusive. Sweeters and Wyeth [28] have developed a model of ‘gameflow’ as a set of salient features for flow experiences in games, and their analysis has subsequently been modified for pervasive games by Jegers [12]. These models are well suited as a basis for post-game interviews and surveys, but require a level of reflection that is difficult to maintain during an ongoing game session without breaking the gameplay.

A related concept is ‘immersion’. Jennet et al. [13] have developed a model for measuring immersion based on the analysis by Brown and Cairns [6]. Their model is closely related to that of cognitive absorption as a state of deep involvement with software [1], and to some extent also to the notion of presence in virtual reality research [23]. But the concept of immersion also faces challenges when recast into a self-reporting tool, as it is basically considered a quantitative measure. The problem is that the more immersed a player becomes; the more difficult it becomes to break off from the game in order to report the current experience. It is likely that players will report less often as they get more immersed into the game, rather than report higher degrees of immersion.

There are however more aspects to immersion than just a scale. Based on an analysis of the vastly different preferences that people show towards digital games, Ermi and Mäyrä [9] provide a candidate model that focuses on qualitative differences between different types of gameplay experiences. Ermi and Mäyrä suggest a multi-dimensional model of immersion consisting of sensory immersion, challenge-based immersion, and imaginative immersion. In their model, sensory immersion is closely related to the concept of presence, whereas challenge-based immersion is closely related to Csikszentmihalyi’s concept of flow. Imaginative immersion has no obvious counterpart in previous literature on immersion.

The model proposed by Ermi and Mäyrä was developed for digital games and their concepts do not immediately transfer to pervasive games. However, a very similar model, the ‘three-fold model of role-play’, was developed by the table-top role-playing community in the early nineties [16]. This model distinguishes between games focussed on story, challenge, and world simulation. Bockman [4] gives the model a slight twist by applying it to the ways in which an individual player can engage in a role-playing game: through addressing a challenge, participating in a story, or pretending to be a person in an imaginary full world. Of these three very similar models, we selected to base our development on Bockman’s model, as it is focused on the individual player’s experience and relevant for a wide range of pervasive games.

3.2 Activity-Related Emotions

Several studies of games (see e.g. [11,18]) aim to capture the players’ emotional responses. These approaches have often been based on Russell’s circumpolar model [22], which maps basic emotions to a two-dimensional matrix where the axis correspond to valence and arousal, respectively. Some approaches use biometrical data [11,18] in this context as a way to capture the experience automatically; however it remains difficult to understand how such measurements to the actual experience which remains fundamentally subjective. The emotional responses to a game are particularly difficult to capture in post-game interviews, as the momentary emotions are particularly prone to be re-evaluated in the light of the full game experience. Hence, we decided to create a way for players to record their momentary emotions while gaming.

Russell’s original model of emotions does not directly transfer to the emotions that players express in relation to activity. Rather than being ‘happy’ or ‘sad’, the play experience is better framed by emotional expressions that are related to activity, such as ‘engaged’, ‘stressed’, or ‘bored’. Most of these emotions still map well into the two-dimensional model, but they also express something about the extent to which the player is in control of the ongoing activity. Although this potentially could be interpreted as a third axis, we decided to stay with the original two-dimensional model but use activity-related terminology to express emotions during our design process.

3.3 Measuring the Ineffable

Experience-oriented HCI researchers emphasise the ineffable nature of experience and experience-related emotions. Boehner et al [5] emphasize that “the production and interpretation of emotion – of national pride, justifiable anger, or shame – is social and cultural in origin”. Boehner et al. advocate an interactional approach to understanding emotions in interaction.

“This approach sees emotions as culturally grounded, dynamically experienced, and to some degree constructed in action and interaction. … an interactional approach moves the focus from helping computers to better understand human emotion to helping people to understand and experience their own emotions. … Systems inspired by the interactional approach to emotion emphasize the expression of emotion in a co-constructed, co-interpreted fashion. Measures of success for such systems therefore do not focus on whether the systems themselves deduce the ‘right’ emotion but
whether the systems encourage awareness of and reflection on emotions in users individually and collectively.”

It is in the context of this tradition that we design our tools as well as the evaluation method. The goal of our tools are not to unambiguously express an activity or an emotion, but to support an ambiguous reporting of what the players currently and immediately feel prone to report. Verbalising or quantifying the experience should not be required. Our approach is inspired by a few predecessors, most notably the Sensory Evaluation Instrument [17] and Emoto [26, 27]; both systems support vague non-verbal reporting of emotion but do not target gameplay experience.

4. DESIGN PROCESS

The design project consisted of three phases. The first iteration included a conceptual pre-study where we tried out the basic concepts for both tools, and developed initial designs for implementation. The second iteration consisted of an in-lab user study after which the system was slightly re-designed. In the final phase we used the tool in the evaluation of the pervasive game Interference [3].

4.1 Design Aim

As described in the background section, we had selected two models for experience that we at the same time considered appropriate for studying pervasive games, and possible for players to report on while engrossed in a game. In line with the interactional approach to experience research [5], we aimed for a reporting tool that would support vague and nonverbal feedback, complemented by a post-game interview in which the recordings would be discussed and reflected upon. This approach to studying game experience is very different from that of quantifying the game experience along a range of pre-defined parameters. The factors that the tools aim to capture are just means to this end; they capture something, but far from everything, of the player experience.

The core design consideration for the tools is thus what players are able to report. In the post-game interview session, the trace of feedback may be complemented by other sources of information, such as information about the players’ location, ingame events, or even biofeedback from sensors worn by players.

4.2 Conceptual Design Iteration

We first sought to find suitable interaction concepts and layouts for expressing the two selected experience models in the player device. One tool would be used to express emotions according to the Russell circumplex model; the other would be based on Bockman’s three-way model of immersion.

Due to the intended usage situation for the tools, we identified quite harsh requirements on the design. The players might find them selves be playing mostly outside, be constantly on the move and involved in activities that might require physical strength as well as their both hands. In order to not disturb the players’ attention on the game, the tool must also be very easy to learn, use (and ignore). Each report should preferably require one single interaction with the device; be available at a brief glance and with immediate feedback provided on the device.

4.2.1 Choice of Player Device and Initial Design Choices

We selected the 3G IPhone as our device. This device offers a comparatively large touch screen as well as almost constant access to data communication. As the IPhone offers touch-based interaction, a natural choice was to let players touch and drag objects on the device screen as the sole model for interaction. We decided to use a commercially available ‘sports armband’ to strap the phone onto the players’ arm, converting it into a constantly available wearable.

The emotion-reporting tool was based on several predecessors used for similar purposes. A commonly used approach in marketing is the usage of SAM, the “Self Assessment Manikin” [21]. In SAM, each emotion is expressed using an iconic face with an emotional expression. Emoto [26,27] instead uses abstract representations such as colours, shapes and movements to express emotions, and from here we picked up the use of colours. In Emoto, the natural color wheel is mapped against Russell’s circumplex model. This mapping is of course ambiguous, and the emotional valuation of colours varies both with culture and person. However, the Emoto studies show that there is a surprising amount of correspondence in how different people interpret the different colours.

The activity-reporting tool had no corresponding predecessors for concepts as the emotion-reporting tool. The three-fold model has been represented by a triangle in literature, however there was no reason to believe that this would suit a self-reporting tool directed towards players.

4.2.2 A Conceptual Pre-Study

The purpose of the initial study was to select interaction concepts and interface layouts for both tools.

4.2.3 Pre-understanding of emotion concepts

In the first part of the pre-study we let participants match facial expressions, a coloured wheel corresponding to the circumplex model, and words expressing emotions to each other to understand their intuitive understanding of these different forms of feedback. The faces were drawn as iconic faces (akin to smileys). As words, we used activity-oriented words for emotions (stressed, bored, and etcetera). Finally, the colour wheel was inspired by the emotion wheel developed by Stål for Emoto [26] but simplified to fit the limited screen space of the IPhone.

The participants commented that the use of faces and colours was an unusual but pleasant way to express emotions. The participants did not completely agree on how to use the colour red, but most used it to express anger. However, we noted that the participants tended to associate light tones to positive emotions and dark tones to negative emotions. As a result of this study we decided to rearrange the colours in the circle, while keeping Russell’s model intact. We also adapted the colours by using a lighter shade in the positive semi-circle than in the negative half. The red-toned colours were placed on the negative side on the wheel, since most of the participants associated red to anger.

4.2.4 Simulated Use of the Emotion Reporting tool

In the second part of the study, we let the players report their emotions by placing iconic facial expressions on the colour wheel. As stimuli we used YouTube clips varying in mood and story content, with a mix of excitement, sadness, surprise and humour. The participants were told to report on the emotions they felt while watching or after watching each clip. Players had little problem reporting their emotions related to the clips in this way,
but they always used more than one face for each clip. To accommodate for this behaviour, we chose to allow users place up to three faces at the same time in the implemented tool.

4.2.5 Simulated Use of the Activity Reporting tool
The pre-study for the activity-reporting tool focussed on establishing a good screen layout. Based on three pre-selected icons to represent the activities (two masks for acting in role, an ear for listening to a story, and two chess pieces for gamist play) we tested three different interface layouts on paper with paperclip dots to be used as indicators (see figure 2).

The participants had no problems in associating the three icons to activities. However, none of the screen layouts was liked by the participants. One of the participants suggested an alternative layout for the triangle, where the three axes were arranged from a common origin in the middle and spanning a circle that encompassed all three. This arrangement made the two interfaces similar as well as made good usage of the device’s screen space. We therefore decided to try it out in the first implemented iteration of the tool.

4.2.6 Prototype Design for the Reporting Tools
The first prototype design of the emotion and activity-reporting tools is shown in figure 3. The design of the emotion-reporting tool allows the user to place up to three ‘dots’ on any place of the circle and move them by point-and-drag. In the emotion-reporting tool, these dots appear as faces that take on the emotional expression corresponding to the reported emotion as a way to provide feedback. The user is free to place the dots anywhere on the circle; on the periphery or towards the centre. The intention is to let players express more mixed or less intense activities and emotions by placing a dot closer to the centre.

Figure 1. Example placement of words and faces on the natural color circle.

Figure 2. The screen layouts tested for the activity-reporting tool

5. Second Iteration: Studying the tools in use
Before the second study, we implemented a fully functional device with the two tool interfaces from figure 3. In this study we put the tools to use in a game session in order to gather user feedback on the functionality, interpretation of and experience in using the tools during a game. Since we intended the study to inform the final design of the tools, the study was small and focused on gathering qualitative feedback from the participants.

In order for us to monitor how the players used the tools during gameplay, the study was placed in a lab setting. We let three players use the tool during a table-top role-playing session. The tools were passed around in the group during the game session, so that all three could test both tools. Before the game session started, the tools were handed out and briefly explained. After the session, we interviewed the players about their usage of each tool. They were also asked to answer a web-based questionnaire where they could provide further feedback on the tools.

In this study we saw that players largely placed their dots on the periphery of the circle, and did not use the screen space towards the middle of the circle to express less intense or more mixed immersions. One player also complained about the lack of means to express immersion. Furthermore, the players never placed more than one ‘dot’ at the same time. Although all three players reported that the tools would disturb their gameplay a bit, they also agreed that this disturbance was acceptable.

“...when you’re acting out, then it might get difficult to stand aside your character, because you're immersed in some extent...it might be disturbing” (interview comment)

The major concern expressed by the players was that the tools were unlikely to be used during the most interesting periods in a game. From observing the players during the game, we also learned that the participants tended to use both tools the most right when they received them. After using them for a while they left them off.

When the evaluator triggered the phone vibrator to encourage players to report, they were in general slow to react. First, they would conclude their ongoing activity, and only then pick up the device. The participants reported that they ignored the vibration when it was sent during intense moments in the game. This confirms our hypothesis that the most interesting thing to note might be the lack of reporting during certain time periods.

5.1.1 Feedback on the Activity Reporting Tool
All three players found the activity-reporting tool easy and relevant to use. A problem with the tool was that the players felt a need to change their state very frequently. The game session required the players to constantly shift levels of activity in short amount of time. Players tended to report on these micro-level changes in gameplay.
5.1.2 Feedback on the Emotion Reporting Tool

The emotion-reporting tool received rather negative feedback. One of the players stopped using it altogether as she could not figure it out, and the other two also reported problems.

The players found the colour circle too difficult to interpret as an emotion palette. Since the faces would appear only as feedback on a placement, searching for the right expression took too much effort and became a disturbance. A related problem was that there were too many similar expressions and that some of the emotional expressions (such as scared or bored) were difficult to identify. One of the experienced players still liked to use the tool in less active periods, as a way to reflect on the game.

5.1.3 Redesign of the Player Reporting Tools

As a result of the study we decided to change the design of the emotion-reporting tool. The details of the interface were re-drawn by a professional graphical designer. We decided to keep the colour circle as a way to provide visual guidance at a glance. However, the background was redesigned to explicitly show Russell’s matrix, with the axis drawn and marked (see figure 4).

We also decided to complement the faces with pop-up texts when moved or tapped. Despite the fact that the participants never placed more than one face, we decided to keep the option to express more complex emotions, as pervasive games may provide players with more time to interact with the tool and the option might still be important for some games.

The activity-reporting tool was kept intact, but we decided on recommending players to use this tool to report the ‘general stance’ of the current game activity, rather than to report micro-level changes in activity. This should make the task of reporting on the tool more manageable and also lead to more relevant report logs for the purpose of post-game interviews.

6. DESIGN OF A VISUALISATION TOOL

The Babylon system consists of the user reporting tool described above, a log server, and a visualisation interface which is used in post-game interviews. The report server and analysis backend is written in Ruby, using the Rails framework. Data storage is done in a standard mysql database. All visualization is written in client-side javascript using the Mootools framework.

6.1 Design considerations

Early on, we decided to build the tool so that it could be used both in runtime while the game was ongoing, and post-game. All traces were thus communicated in real time from the device to the server and immediately visualised.

As discussed previously, post-game interviews benefit from reviewing other data than just the trace of player reports. In the implemented tool we choose to include a GPS trace from players, based on the fact that the 3G IPhone contains a GPS. Further more, earlier studies (see e.g. [26]) have shown that notes tied to important events during the game provide an immediate context for the game-masters when revisiting the data. Therefore, we decided to include a bookmark system, where the game-master or evaluator could, at an important time in the game, enter a note. This system allows the evaluators to make arbitrary notes about important events in the game, irrespectively of the source of information.

As mentioned previously, our hypothesis was that players events in the game, irrespectively would be less likely to interact with any provided tools while engrossed in engaging gameplay. This made it important to visualize changes in interaction for the evaluators, as this could reveal important information about the current state of the game and the player’s emotional attitude towards the game.

6.2 The Babylon visualisation tool

The visualisation interface went through several iterations in collaboration with the evaluators. The first iterations were more or less sketches - providing the game master with the reported data and the current location of the players (i.e. the devices carried by the players). These where found wanting in the aspect of providing information regarding downtime and pauses in the reporting. Different concepts where tested, among them a feed-based interface and a timeline-based interface. In the end the Simile time-line\(^1\) was selected as the primary means of interacting with the log data. The continuous updates of the timeline made it clear to the game masters when the players were reporting and when not. Clicking in the timeline implemented a kind of ‘time-travel’ and would immediately display the complete state of data at that time. The final layout of the visualization tool is shown in Figure 5. It is inspired by a multitude of sources, among them the

---

1 The Simile Timeline is an open source Project from MIT. It implements a DHTML-based widget that maps events on different timescales. See http://www.simile-widgets.org/.

---

Figure 4. Second design iteration of the player reporting tools.

Figure 5. The Babylon visualization tool.
Exhibit platform (see MIT Simile project). During play, the interface is updated in real time, showing the current position of the players and the current status of the two reporting tools in play. To navigate the time-stamped log, the evaluator selects a previous report icon or time-stamped event from the timeline. The status of the reporting tools and the players’ position at that time are then shown above. The visualization can also be toggled to instead display a large, real-time updating map of the entire game area showing the current position of the players and a fading trace of their movements for the past hours. Finally, the system supports pushing notifications to the player carrying the devices, in the form of a vibration - something that was done through the visualization-interface.

7. EVALUATION

In order to test the tools in use, we used them as part of the evaluation of the pervasive game Interference [3]. This game is well suited for testing the tools, as the game is restaged and played about the same way in every setup. All players also play together as a single group throughout the game, although they have different responsibilities within the group. Interference is an adventure game in which a group of six to eight players take on the roles of technicians, tasked with finding the cause of a large Internet breakdown. The storyline changes during the game, and eventually the players realize that they have gotten involved in a bizarre family conflict with no easy solution.

For this study we were particularly interested in understanding how players would react to using the evaluation tools in their target usage situation, in a pervasive game played on the streets. We also wanted to understand if the data gathered during a play session could be used during post-game interviews. For this study, we had four hypotheses:

- In line with the role-play study, we assumed that participants would find the activity reporting tool easier to use than the emotion reporting tool.
- The re-design of the interface for the emotional-reporting tool should however make it acceptable to players.
- The participants would use both tools less during intense moments in the game.
- The participants would report more frequently during this study than in the table-top setting, due to the shifts between low and high pace in the game.

7.1 Study setup

The game was staged on three occasions during February 2009 in Kista, Sweden. Six to eight persons participated at every occasion. On each occasion, two people used the tools: one used the emotion and one the activity-reporting tool. In total, 21 players participated in the study, of which six used the reporting tools.

The players were introduced to the tools as part of the introduction to the game. After the game, the players were interviewed about their experiences, partly based on the Babylon logs. They were also interviewed about their experience with using the tools. Finally, players were asked to fill in a web-based questionnaire which contained questions related to the tools. These questions were identical to those posed in the questionnaire in the second design iteration.

7.2 Results

According to the survey responses, both tools were considered somewhat frustrating to use (3 of 6 players). The interview comments indicate that this frustration was primarily due to hardware issues: players would report that they found it difficult to interact with the touch screen through the plastic cover (or due to cold) and that the special gloves used for this purpose did not work well. In general the hardware issues made them less prone to be precise in their reports.

“Sometimes it didn’t work. Sometimes you got something, and it’s just “that is just OK”

“In the end you got so frustrated that you just did something, just placed it somewhere where it was like good”

(Interview comments)

Compared to the tabletop study, we noted some differences in how players placed their markers when reporting on the emotion-reporting tool. This time players used the entire circle when reporting. This is most likely a result of making the axis more explicit. There was however only one player that reported to have tried using a ‘mix’ of emotions by placing several markers on the circle at the same time. This player also reported that he did so only when he felt that they were confused or bored. This might potentially indicate that the more varied pace in gameplay that a pervasive game offers, the more time is given for reflection.

The reporting logs indicated that players reported less often during intense moments in the game, and the players also commented on this during the interviews.

“And if it’s really fun and good in the game, then you forget to do updates, because then you don’t think of things that are out of game”

“When things did happen, then I forgot to use it. It was only when I didn’t have much to do”

(Interview comments)

However, players would often remember to report their emotions after intense periods. These delayed reports were often based on what the players had experienced during the intense moments, as a kind of reflection. Again, this indicates that the varied pace of a pervasive game is well suited for this type of reporting tool.

The reminder vibrations were not even noticed in this setup, as no player noted the vibrations through their thick winter clothing.

7.2.1 Activity-Reporting Tool

One player was neutral and two positive towards the statement ‘The self-reporting tool provided a good way to report the types of activities carried out within the game’ (grades 4 and 5 on a scale where 0 corresponded to ‘completely disagree’ and 7 to ‘completely agree’). Two out of three players also stated that the icons were well chosen and easy to relate to the current activity in the game, and the tool was considered easy to learn and understand.

In the interviews, two participants expressed some concerns about whether they had reported ‘correctly’.

“There were three different things, I did my interpretation of it. I’m not sure though if it was right”

“The ear, is it that we are listening to a story or when you’re (referring to another player) talking so much in the phone [with the game characters]. Is it gaming we are doing or now, or listening?”

(Interview comments)
As the purpose of the tool was to allow focussed post-game interviews, there was no ‘correct’ way of reporting an activity. In fact, the post-game discussion made it clear that the mere act of reporting made players create a strategy for how to reflect on their own game activities. The participants had no problem in remembering large parts of the game and how they felt and acted at certain moments in the game. However, reviewing the log from the visualization tool helped them to remember why they reported as they did and how they had interpreted the details of the interface on the player reporting tools.

There were some differences in how players reported on the same activity. One player explained that she placed the marker on the ear when they were listening to messages as well as when talking to the actors (characters in the game), while another player reported that he placed the marker on the theatre masks when they talked to the actors. The difference might be due to how the players engaged in the game. The first player participated in a group that addressed the game in a rather gamist way, focussing on solving the mystery and using the actors as a source of clues. The second player participated in a group that engaged deeply in the story and adopted personal roles in communicating with the actors.

7.2.2 Emotion-reporting Tool

All players reported themselves as neutral towards the statement ‘The self-reporting tool provided a good way to report the types of emotions experienced during the game’ (game 4 on a scale where 0 corresponded to ‘completely disagree’ and 7 to ‘completely agree’). However, this tool received more criticism than the activity-reporting tool. Two out of three reported that they appreciated the colour wheel, but they also reported that they were not able to relate to the facial expressions with what they experienced in-game.

There was some criticism directed towards the feedback supplied by the tool, in particular the textual feedback on the reported emotion. One user commented that this interfered with the possibility of creating an own interpretation of the tool. This respondent suggested removing the labels whilst keeping the colors and axes. This is consistent with the interactional approach to emotions: verbalisation is not necessarily helping. The facial expressions used as feedback worked better but were not ideal: some players reported that they sometimes were surprised by the facial feedback. The problem was primarily due to the fact that each face corresponded to a fairly large portion of the circle.

Two users liked the tool and found it easy to use. One particularly liked the way it helped him to reflect on the game while it was running. An interesting aspect of this was that these players both believed that the game masters were using their reports in managing the game. Whenever they felt confused and reported this and something new then happened in the game, they thought it was the game masters reacting to their reports.

‘…it can be a good way to know when to portion out extra bits of information or help if we think, Okey now we’re stuck. Aa, okey, then we’ll give them [the players] this information package so that they continue to the next step. And yes, to use it like that, I think is a good way…’

(Interview comment)

The third user reacted strongly against the tool. His main criticism concerned that the very purpose of the tool, reporting on in-game emotions, was too disturbing and affected his game experience. He argued that the very act of “stepping back” from the experience to give a subjective response was too time consuming and affected his ability to immerse in the game. He also felt being restricted in his reporting, as the labels for the facial expressions were too extreme in comparison to what he wanted to report.

7.2.3 Visualisation Tool

The Babylon visualization tool became more than a tool for just the post-game interviews. Also, the game masters kept an eye of the visualisation throughout the staging of Interference. Although we have no documentation of this, it is possible that this also influenced their way of staging the game just as some players had inferred.

During the post-game interviews, the evaluators used the tool to spot interesting moments during the game, as well as interesting trends in how a player had been reporting throughout the game. We could also note that the players tended to remember (or mention) more about their experience when probed by the evaluator. An example of this was when the evaluator asked a player why he seldom had reported in the ‘role-play’ part of the activity tool. His response was that he seldom felt that they had been role-playing in his group (a statement that the other participants confirmed).

8. CONCLUSIONS

The goal of letting players self-report is to gather immediate first-hand data on their subjective experience. From our study, we see that it is possible to use self-reporting tools to report on both activities and emotion during the course of a pervasive game. Furthermore, an interactional approach to the design of such tools has several advantages, in that it lets players infuse their own meaning into the reports and to focus on the game during its most intense moments. In the interactional approach, players report in a vague manner and the logs are primarily used as a background for discussion in post-game interviews.

The players in our study were also surprisingly consistent in their use of the tools. The differences in how the tools were used seem to have corresponded to actual differences in how players engaged in the game. We also noted that game masters kept watching the tool usage during the staging of the game. This indicates that the tool reports give useful information for game mastering, and that in the long run they may even be used as stand-alone tools for evaluation. The latter would require gathering a corpus of log data from a large number of game sessions with different games, and cross-referenced with quantitative data from post-game interviews.

Reporting on activity seems to be easier than reporting on emotion. Players reported that they found it difficult to report on emotions, and it also proved harder to design a good interface for this tool. The players expressed this as a problem of distancing; in order to report on your emotion you have to distance yourself from the game to think about how you feel. By contrast, reporting on activity seems to be more immediately accessible.

It is impossible to develop self-reporting tools so that they never disturb players. However, the interactional approach proved to be beneficial in this respect, as it made players feel less obliged to report regularly and accurately; thus focussing on gameplay whenever the game demanded so. In general, our hypothesis that players would report less often during intense periods of the game was confirmed. Still, we found that players were prone to report also on intense moments of the game, only slightly delayed. Since
pervasive games typically vary greatly in pace, this shows that self-reporting is a viable option for these games.

8. ACKNOWLEDGMENTS
This work was carried out as part of the Mobile Life excellence centre. The authors wish to thank the anonymous players who took part in the interference study.

9. REFERENCES