

Towards the Wild: Evaluating museum installations in semi-realistic situations

Eva Hornecker and Emma Nicol

Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK
Eva.Hornecker@cis.strath.ac.uk, Emma.Nicol@cis.strath.ac.uk

Abstract

In this paper we discuss our experiences of conducting user evaluations of technological installations for the new Robert Burns Birthplace museum and the challenges associating with doing this. We believed it to be important to avoid a sterile lab situation, and to be able to observe the social dynamics that evolve around museum installations. As we were unable to conduct studies in the museum itself we devised ways of taking them out of the lab by basing them in a more museum-like or semi-realistic setting. We took efforts to recreate the installation setup when this was likely to influence interaction patterns and could reveal usability issues. Moreover, we invited young families and adult groups to take part in the evaluation sessions to investigate how they would interact with these as a group. We explain our motivations and the tradeoffs involved. We focus on discussing the effects that the semi-realistic setting may have had on the user feedback obtained, reflect on the differing user behaviours that might emerge in the genuine museum setting and make recommendations for future studies.

1 Introduction

In the museum and heritage domain, projects are often on a very short schedule, with limited sources of funding. Most technology-based installations are being developed by SME art and design companies, who lack a budget for in-house research and evaluation. This has several consequences. Installation designers primarily rely on their intuition and design experience, and design companies rarely receive detailed feedback. Even imaginative, engaging and visually interesting installations may suffer from usability issues, as installations rarely undergo user testing before deployment. Observations of visitor interaction in museums (Hornecker and Stifter, 2006; Hornecker 2008) found that minute details can hamper the user experience and prevent people from appreciating what an installation has to offer. Moreover, there is not much literature about how to best go about conducting formative user evaluations of museum installations.

Most of the museum literature focuses on describing approaches for and findings of visitor studies within museums (Macdonald 2008). These tend to investigate aspects such as navigation, dwelling times, the effects of room layout and exhibit location, and visitor learning. We have found little literature describing the development process of installations and in particular, how to go about testing early prototypes which are not yet at a stage where they could be temporarily deployed inside the future setting. As described above, even with the best content and concepts, glitches in usability can hamper the use experience, and may even prevent what Allen (2004) refers to as ‘immediate apprehendability’ – visitors understanding within the first few seconds of approaching an installation how to interact, and getting the gist of what it is about and what is the reward for interaction. Formative evaluations and usability tests aim to

identify problems early-on in development by observing five to ten participants using a system, providing recommendations how to improve it. These methods are in wide use in the software industry where it is important to optimize website layout and navigation, and to ensure that interface metaphors are well understood by users.

Traditional usability testing focuses on efficiency, speed, and error rates in individually performed tasks. This assumes core tasks which most users will want to achieve (e.g. place an order in a webshop). In a museum setting it is difficult to tell what ‘a task’ might be – visitors primarily want to be entertained and educated, and their aim might change with what a system offers. Free-choice learning is intrinsically personal (Falk, Dierking and Adams 2006), and thus cannot be prescribed through tasks or measured in learning outcomes. The museum setting offers many distractions, and an installation that is not immediately easy to use and satisfying is quickly discarded for other attractions (Allen, 2004). User testing of museum installations thus needs to investigate both usability and enjoyment. Moreover, the museum situation is inherently social – visitors come in groups, want to share an experience, spend time with children or friends (Kelly et al 2004, Heath and vom Lehn 2002), and are indirectly affected by the presence of other visitors (vom Lehn et al 2007, Hornecker 2010). The social context is identified as one of the core aspects of the visitor experience in the literature (Falk, Dierking and Adams 2006) In this sense, museum installations are more akin to e.g. social/party games, which often are tested by handing them to beta users, or playtested in labs that simulate a living-room (Pemberton and Griffiths, 2003; Isbister, Schaffer 2008). Study of team-play has proven useful for game designers, providing insights into how teams share control and communicate.

In this paper we discuss our motivations and experiences in evaluating a set of interactive installations for the new Robert Burns Birthplace Museum (short RBB museum) with potential users. Our team was responsible for running formative evaluation studies of prototypes for four touchscreen and one table-based installation and providing reports to the National Trust Scotland and the design company developing the systems. Our reports provided feedback on interaction problems, how users enjoyed the installations, whether they understand how these relate to Robert Burns, and suggestions for re-design.

The main visitor groups for the RBB Museum are young families and so-called ‘empty-nesters’. Most of the installations are aimed at engaging children with the themes of the museum: Burns’ life, his poetry, and the era he lived in. They have the form of mini-games, for example, creating a photo-fit of what Burns might have looked like, and having a mouse jump up to grab cheese in the rhythm of a poem.

To achieve more ecological validity, we decided to test with groups of users, and to invite families to ‘play’ with the prototypes. One of our main motivations was that interaction issues experienced by a group might be very different from those discovered in single-user trials. Research in Computer Supported Cooperative Work (CSCW) has revealed how requirements for group work can sometimes conflict with requirements for individual work (Gutwin and Greenberg 1998). This very likely would also be the case with edutainment activities. The ability of a game to entertain a family might be essential

for its overall success. Would siblings be able to play the games together or would they be fighting? Would the games be interesting enough for children of different ages and their parents to stand back and observe while others play? What age group would like the installations? Might the presence and help of parents make the games accessible to younger children, who are not able to play them on their own? Some of the recent museum studies literature has highlighted the role of family and parent-child interactions, with parents explaining, pointing out things, asking questions (Kelly et al 2004, Sanford et al 2007). This would mean that having children use these systems on their own is not representative of the kinds of interactions occurring in the museum setting. Moreover, one of the installations, an interactive table, was designed explicitly for cooperative (and competitive) gaming.

In this paper, we describe how we aimed to account for these factors and discuss some of the questions this left open. While we tried to emulate the group situation of a visit, our user trials nevertheless took place in an artificial situation: families were invited, remained together for the study, with no other activities competing for their attention. The museum was being rebuilt at the time, and we were thus unable to run studies in the real location. Moreover, the versions we were provided with were early prototypes. Our study helped to uncover many issues that could be resolved for the final versions of the installations, including issues relating to group coordination and play. But we anticipate that there might be issues specific to the museum situation that are hard to replicate in an evaluation trial setup.

The aim of this paper is to contribute to a reflection on ways of user-testing prototypes aimed at group and family use, and the potential limitations of doing so in a semi-realistic setup. Most of the detailed findings of our evaluation sessions are not relevant in this context, being specific to the installations and the flaws of early prototypes. Our summary of findings thus focuses on issues related to the study setup and our simulation of a group use situation. We present a summary of issues, reviewing our observations with the question of whether similar behaviours would occur in the museum. As follow-up research, we currently engage in an ethnographic-style observational study to compare behavioural patterns from the user trials with the 'wild life' in the RBB museum.

2 Context and Study Setup

The Robert Burns Birthplace Museum in Alloway, Ayrshire comprises the cottage where Scotland's greatest poet was born, some of the landmarks providing the setting for his works, and a recently enlarged and rebuilt museum. The new museum opened in late 2010. It is a popular day trip location for young families, school fieldtrips, and senior adults. One of the aims of the rebuilding of the museum was to engage young children (and adults) in a playful way with Burns' life and poetry. The museum makes great efforts to be family-friendly, offering special activities for children, and includes a number of interactive multimedia installations.

Burns is lived heritage in Scotland (and to some extent in the wider UK), where a majority of people will learn some of his poems in school and many will at some point

attend a Burns Supper. The RBB museum displays objects for veneration (original letters and objects), but in multiple ways also invites curiosity (some artefacts are in cabinets which visitors are invited to open) and playful interaction. Children are allowed to run around and play, and besides the multimedia installations, there are a number of non-software interactive stations, where children may for example draw. During our recently started initial ethnographic observations in the museum we observed that this playfulness extends to many of the older visitors.

Our team was commissioned to conduct formative evaluation studies of five of these installations in summer 2010. The software was being developed by Spiral (London, UK). We were provided with one installation at a time, and would be handed the next installation prototype after submitting our report within a given time window. Our remit was to focus on usability, overall enjoyment and game play issues, and to provide concrete recommendations for how to improve the systems. Given the museum was still being built at this time, it was clear that questions of visitor learning would be hard to address, as the systems would be experienced out of context. Moreover, the installations on their own do not have learning objectives in a traditional sense. They are part of a ‘show not tell’ interpretation strategy which is to help visitors enjoy Burns’s heritage in an imaginative and interactive way, and are particularly aimed at engaging young children in a playful way with themes of the museum. We were thus only asked to investigate whether users would realize what the installations were ‘about’ and whether they felt they had learned anything.

We now briefly describe the installations and then focus on how we set up and ran the user study sessions.

2.1 Overview of the Installations

We here focus on four installations, aimed at children and teenagers, most of which were touchscreen-based. For the purposes of this paper, we only provide a rough overview of these mini-games. It should be noted that at the time of testing, some of these were incomplete, sometimes with only a subset of content being implemented, in other cases employing simple line drawings.

PhotoFit (see figure 1 right) presents the visitor with contemporaries’ descriptions of Burns’ appearance (“he had fiery eyes”), who then create a mugshot, selecting facial features from a series of images. *PhotoFit* is purely touchscreen-based. One can select the nose, mouth, left and right eye (individually) and the overall face, including hairstyle. Touching either on the facial features or the headings on the top navigates between different sets of images and brings up historical descriptions of how Burns looked like. In the museum, this installation is now set next to portraits of Burns’ family.

The *Poetry game* plays the poem ‘To a mouse’ and a piece of Tam O’Shanter spoken by an actor, while the poem’s text moves over the screen. An animation of a mouse runs beneath the text, and at the moments of emphasis a piece of cheese is placed over the words. The user has to press a button in the rhythm of the poem, whenever the mouse is

under the cheese, to make it jump up and eat. After completion, one can read the entire poem and is presented with a score of how many pieces one managed to catch. This installation aims to create awareness that poetry has rhythm; the poem is usually learned in school. This installation uses a mix of interaction technologies. While the game is started and one can choose between the two poems by touching the screen, a large button under the screen is used to pound the rhythm of the poem.

The *Spooky Stories* touchscreen game illustrates the poem 'Tam O'Shanter'. In the poem, Tam, a drunken villager leaves the pub on a dark night, encounters mysterious creatures on his way home, running across the graveyard, until he jumps on a horse and rides over the bridge, where the spooky creatures cannot follow. The installation shows the village, and has the user press the creatures from a 'virtual transparency' onto the background and move them into the scenery (figure 1 left). The user can navigate through the scenery by moving the transparency to the side of the screen or pressing the arrows on the right and left of the screen. Once completed, an animation shows Tam going home past the creatures.

Finally, the *Burns Supper Table* is a multiplayer game in five rounds based loosely on the tradition of the Burns supper (see figure 2). It is played around a table, with a projection from the top. Each player has a physical button to interact; the table itself is not touch-sensitive. First, players have to press their button when their segment of the table is highlighted, showing it is their turn, to play the notes of a well-known tune (played during the 'bringing in of the haggis' during a Burns supper). In the next round, players stab haggis (a kind of sausage) that move around the space, the button press making an image of a knife poke out. Then, a memory card game starts, alluding to the tradition of giving an 'immortal memory' speech on Burns. Then, reminiscent of the 'toast to the lassies and laddies', players can throw insults and compliments (in the Scots language) towards two figures in the centre. The game ends with the song 'Auld Lang Syne', and players need to make a circle of figures swing their arms in time with the song by pressing the buttons. If they manage to keep in time, their segment lights up.



Fig 1. Tam O'Shanter (*Spooky Stories*) Touchscreen with transparency on top of the scenery showing the village. PhotoFit game with only the hair being selected so far, and the rest only being indicated by sketch drawings. The quotes describe Burns' hairstyle and colour.

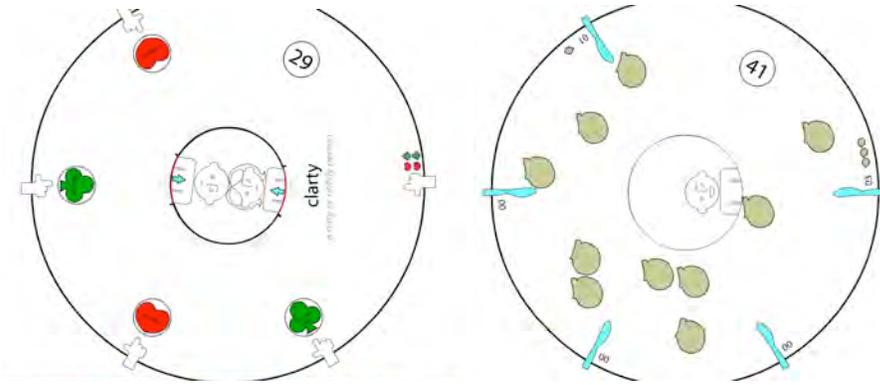


Fig. 2. Some of the games on the Burns Supper Table. Left: One player has just managed to shoot a compliment to the girl in the middle (the figures rotate around the centre, making aiming tricky) and the player is rewarded with a translation of the Scots word 'clarty' (filthy, mucky). Right: Stabbing the haggis, with the 'host' in the middle. The buttons are located on an outer rim just beneath the hands respectively knives in the image.

2.2 Setup and Process of the user trials

Since the technical and contextual setup can greatly influence usability and social experience of play, changing how visitors use exhibits (Isbister 2010, Monahan 2005), we took care to emulate the final setup as much as possible and to the best of our knowledge. It was crucial to utilize touchscreens for the tests, as touch interaction has different usability requirements than mouse use and offers more affordances for collaborative action. Moreover, display angle and size influence how easy it is to select targets on a touchscreen and affect how easily a parent can observe and scaffold a child's interactions. The touchscreen games were tested on a large touchscreen (20") to approximate the final installations. We set them up at about the height and angle indicated by architectural sketches. We also built a table of comparable size and setup to the Burns supper table, enabling us to assess how the game was experienced when played in a group and using buttons (see figure 4). For this, we installed five push-buttons in a self-built table and connected these via an Arduino controller with the Flash-based game, which was top-projected onto the table.

Given the RBB museum is visited by many young families and by groups of older adults, it is important that the installations are enjoyable for groups. Each game was evaluated with six to eight families with children of different ages, and by two groups of 'empty-nester' aged adults. Each group had two to four members, sometimes with both parents present, and for about half of the trials, there were two or more siblings. For Spooky Stories and Burns Supper, we asked adults to bring a friend with a child, resulting in a few larger groups (with five to six members). We received written consent either from a caregiver or the participant himself. Each participant was rewarded with an NTS voucher.

In a further attempt to make the setting more natural, increasing ecological validity, we looked for opportunities to conduct the studies off University grounds, and were granted access to the library space within the Kelvingrove Art Gallery and Museum (see figure

3). Our participants approached our study through this well known museum, and many would spend more time there either before or after the session. Our intention for this was to put participants at ease and to have them experience the installation in a museum setting. This location furthermore enabled us to recruit a few participants on the spot. However, recruitment of most participants relied on networking and snowballing via an initial group of parents. Due to the complex technological setup required for the Burns Supper table, it had to be tested in our University lab.

Within the museum library space, we cordoned off a corner between some bookshelves so that other visitors would not disturb us. At University, where we tested the table, we used our lab which at this point in time was relatively empty (see figure 4).

We began the sessions by welcoming participants, explaining the study aims, and handing out consent forms. We then invited them to play while we observed. In usability studies of games (Isbister and Schaffer 2008) it has proven beneficial not to provide task lists, but simply let groups play with little intervention, as it allows participants to focus on the game and forget about being observed. Our only instruction thus was to ‘go and play with it’. Only if the group struggled for a considerable time to e.g. discover the basic interaction mechanism did we intervene in order to minimize frustration and to enable the group to experience the main part of the game – our remit after all was to generate feedback on all aspects of the installations. Towards the end we asked parents if they would like to have a go in case they had not yet played. During all this time, we took extensive notes of our observations, including group interaction style, the interactions between children, parental interventions, any issues with the installations, and also noted down things that worked well and were successful.



Fig 3. Trial setup in the museum library, with researcher observing and taking notes. Children playing with the ‘Spooky Stories’ touchscreen.

After ending the active game play, we asked younger children to draw a picture of themselves playing the game, while parents and older children filled out a questionnaire. This asked them to rate how easy they found the games and how much they enjoyed them. They were further asked to describe what would positively or negatively influence their inclination to use the installation in a museum, to list what exactly they enjoyed and what they found frustrating, what their children liked or disliked, how the installation “could be improved to be better to use and more fun”, and what they ‘learned’ from doing the activity. Sessions ended with a short open-ended group interview, using the

younger children's drawings to get them talking about their experience. In this final conversation, parents would usually provide more context to their questionnaire answers, and sometimes developed further ideas on how to enhance the installations. Each session lasted 40 to 60 minutes.

3 Families in the lab versus 'in the wild' in the museum

Overall, we could identify many issues in usability and collated feedback on how to improve and expand the mini-games. Our combination of observation, questionnaires, and open discussion proved very useful. We found that often questionnaires reflected users' final impressions (of eg. how easy a game was to play), but did not always capture initial problems in figuring out what they were required to do in order to play. In a setting with more distractions, the same group of users very likely would give up and leave (cf. immediate apprehendability, Allen 2004). Some interaction issues, such as difficulties in selecting small touch targets, are rarely mentioned, and users may not be consciously aware of these despite obviously struggling. Open discussions and open-ended questionnaire questions provided a wealth of insights into what kinds of things people enjoyed, and how they could imagine extending the installations.

In the following, we focus on issues related to family group interactions, and on questions that our study setup did not allow us to answer. While we took care to create a relaxed setting for families, and to remind participants that the installations would be within a museum, we were unable to recreate *all* aspects of a real-life museum setting. This means that there may be interaction issues which we were unable to test and predict. To investigate this, we are currently conducting an ethnographic-style observational study within the museum.

The following discussion is informed by our experience from observational visitor studies within museums and a reflection on whether some of our observations in this user study might be an artefact of the study setup. We start by discussing the variety of group interaction styles observed, and how different games evoked different patterns of group interaction. We then focus on the role of adult scaffolding of children's interaction and facilitation of group interaction, mediating between siblings, for example. Finally, we discuss issues that highlight the limitations of our study setup. For example, we could not emulate the interaction between groups of strangers, and found that participants often made suggestions for improvement and extension based on experiences with video games used at home that were not appropriate in a museum setting. We conclude with a reflection on the utility of our attempts to emulate the 'real' museum situation and the tradeoffs involved.

3.1 Different Installations Evoke Different Group Interaction Styles

The style of social interaction differed markedly across installations with a clear pattern emerging for some of the games. Without inviting family groups and extended families (bringing a friend) to our studies, we would have been unable to observe this. In

particular, some behavioral patterns only occur in larger groups with e.g. several children. One of our study aims had been to investigate whether the installations would be fun to be played with a sibling present, or whether they would exclude them from participating, potentially resulting in conflict. We also found that some of the games gave more room for parental involvement than others.

Adults often kept in the back during the trials, and first had their children try. With younger children, the parents often played first to show how to do it. While adults often took part in games that could be played cooperatively by taking turns (especially PhotoFit), parents largely kept in the background for the Poetry game, and only played after their children had done so several times. This is partially explained by the kind of activity the Poetry game entails, which is difficult to share and requires a lot of concentration in listening, watching the screen (the mouse has to be just under the cheese when the button is pressed), and beating the rhythm. When parents or another sibling played this, it provided an additional learning experience for children, as these could now focus on listening to the poem.

The term *scaffolding* covers helpful guidance and assistance that enables a child to do things he/she cannot yet do fully on his/her own. Parents may scaffold by, for example, explaining what to do next ('press this'), reading out instructions, and guiding children's hands. While families engaged in a lot of turn-taking and scaffolding each other with the PhotoFit, the Poetry game was very much an individual activity, with the family watching. Apart from helping to read, there was little scaffolding from parents. This is probably because the game is very fast, leaving no time to explain or give instructions. Here, family members were often competing against each other in rounds, shouting out and memorizing their final score and comparing against previous scores. With Spooky Stories, smaller children sometimes competed and fought for control, even moving each others' hands away. This was almost never seen with the other installations.



Fig 4. Different family configurations around the Burns Supper Table: Two families, a group of young teenage friends with a mother, and two mothers (standing back) with daughters.

At the Burns Supper Table we saw the biggest variation in parental scaffolding, with some parents taking a 'hands-off' approach, leaving it to their children to figure things out. This usually resulted in a less engaged experience, with young children (below reading age) often not understanding what to do. They would press buttons at random times, but not experience the actual gameplay. In reverse, in groups with older children these often were first to figure out how to play and show or explain to their parents. Different to other installations, here parents could play along, and some (who joined the game) clearly enjoyed it, confessing to getting quite competitive. Some families got into competitive mode, shouting out and comparing scores after games. But many also

enjoyed playing as a team, encouraging the others to play correctly (e.g. with the Auld Lang Syne game saying ‘lets see if we can get it to light up’). The Supper table overall seemed to support both team play and competitive play, some of the mini-games being more competitive than others, but leaving the style of play open to participants. Initial observations of the final installation in the museum reveal similar patterns of people explaining to each other what to do and playing enthusiastically against or with others, indicating that our attempts at recreating (or rather pre-creating) the social situation of the museum setting were successful.

3.2 The Role of Adult Facilitation and Scaffolding

As discussed above, parental supervision and scaffolding is often instrumental in helping children use an installation effectively. Parents will furthermore often add context, point out things the child should notice, and engage the child in conversations that relate the current object to previous experiences (Kelly et al 2004, Sanford et al 2007). Moreover, parents are busy facilitating and moderating interaction between groups of children, in particular if these are siblings, trying to minimize conflict and to ensure that all children get their share.

We observed parents reading out aloud for younger children and facilitating sibling interactions. PhotoFit, for example, was highly enjoyed by very young children due to being easy to manipulate, despite considerable amounts of textual description. Even a two year old successfully interacted with it. Parents often tried to orient children towards the educational goal of the installation (emulating Burns’ facial features), and to make them read and understand the descriptions. Parents’ behaviour was carefully modulated reflecting the children’s age and reading ability, reading out aloud, encouraging them to read, or asking questions about what certain phrases meant. Parents further moderated and facilitated turn-taking by siblings.

Parental facilitation and scaffolding was also very important for the Poetry game. Here, parents encouraged children to listen to the poem, and, using the pause when it is shown full-screen, asked whether they remembered it from school, or knew what certain words meant. For this installation, without parental intervention, children mainly saw it as a quick-reaction game, almost ignoring the poem.

The role of parental guidance observed was encouraging in terms of showing how family interaction can make installations accessible and enjoyable for youngsters who would not be able to do so by themselves. The entry age for most installations turned out to be much lower than anticipated by system designers and ourselves. Also, parents regularly used the opportunity to educate their children, explaining the content, and asking children to remember things. This highlights the importance of setting installations up in a way that provides space for parents and allows them to see what their children are seeing and doing (cf. Hornecker 2010), which in turn enables them to comment, intervene and help.

The question left open is whether parents will be as closely engaged in the actual museum situation, or whether they will prefer to have their children entertained with the

installations while they browse the showcases of Burns' letters and belongings, read information panels, and talk with other adults. Parents in our study very likely felt that they needed to supervise their children more closely than they would do without the presence of an observer. The amount of parental engagement may also be influenced by subtle differences in the physical setup, for example the exact height of touchscreens, the floor (many parents sat or kneeled down), and the space available. We anticipate that in the actual museum situation, there will be an even greater variety of group behaviours, including children that are temporarily on their own while adults investigate display cases.

3.3 Fixed Groups in the Lab – More Fluid Constellations in the Wild

In our sessions where we invited families, groups came and stayed together, with no distractions that might result from the late arrival of a group member. Our study setup cannot uncover potential interaction issues that arise from more fluid and shifting activity patterns. Previous studies on interactive surfaces in public settings have discovered that the assumption of one activity being done by a stable group can result in system designs that do not fit with real life situations of groups being in constant flux (Marshall et al 2011), preventing newly arriving members from joining. Moreover, in a museum setting it is quite common for an activity not to be completed, and the next visitor to pick up from this point. This potentially results in interaction problems if the new arrival does not see the initial 'entry screen' which introduces the goal of the game and explains how to interact. For many installations, instruction bubbles (help text) are only visible initially, and disappear after the first touch. In addition, having only one family group at a time for our trials, it was not possible to investigate how easy it would be for additional people to join an existing group. We also could not determine the maximum number of people who can interact productively with the installations.

We observed that children, teenagers and young adults were likely to use several buttons on the Burns Supper Table at once, working two segments (meant to be for one participant each) or even trying to cover three and more free segments (see figure 4, second from left) if there were not enough players to use all of the buttons. For our current observational studies within the museum, we are interested in whether the game encourages strangers to play together. The design of the games should encourage this, as they can be played without having to discuss and negotiate. Players can join while a game is running by pressing a button (this 'un-greys' their segment). But it remains a question whether the tendency to occupy several neighbouring sections of the table might hinder other people from joining, either preventing them from participating or making them hesitate to approach. In our study this did not pose a problem, as group size was fixed, and some parents did not participate. We are also interested to see how different groups or individuals might interact or follow each other at the other installations.

3.4 The Study Setup Influences User Perceptions and Suggestions

Testing in a lab setting, with the prototype clearly running on a PC (not physically embedded and affixed), seemed to result in participants perceiving these as standalone

systems. Even the library setting within the museum did decontextualize. Participants often compared the games with video games that they might play at home, and this seemed to influence their suggestions for extensions and improvements.

Frequently, changes to the installations were suggested that involved ratings (of user performance), competitive versions (played in turns) or playing against a highscore. Children often suggested changes such as difficulty levels that would prolong the activity or make the games very complex. Competitive games can perform well in a museum setting if they allow simultaneous gaming. Having users take turns playing against each other does not work as well, as it prolongs the activity, and can make it difficult for new users to join. Highscores would further complicate the setup. These suggestions are sensible for home video games, but can be detrimental in a museum context where other visitors will need to wait and might be frustrated if they do not get to have a go within reasonable time.

We saw that children were determined to complete every bit of the activities, for example in the Spooky Stories to take all figures off the transparencies and place them in the scenery. Even though children liked all of the figures, we recommended to reduce their number to ensure a better throughput. For the Poetry game, where users had to tap in rhythm to the poem being played out, many participants suggested giving penalties for tapping at the wrong time. But for games aimed at engaging different age groups, it could be discouraging to tell the user how well they performed. We saw that very young children could enjoy interaction even if they did not perform well, and older users often took several rounds to figure out what the game asked them to do.

Adults often commented that it was hard to connect the games with Robert Burns, and thought that anybody who e.g. did not know the Tam O'Shanter poem well would not recognize it in the Spooky Stories installation. This, again, might be very different in a museum setup which physically contextualizes the installation through its surroundings and, for example, enables parents to draw children's attention to how the installation relates to surrounding exhibits (cf. Hornecker 2010). This might provide additional learning opportunities that cannot be replicated easily in a lab study – especially if the final layout and arrangement of museum pieces, posters, labels etc. is unknown at this stage.

Similarly, participant feedback on aesthetic aspects needs to be taken with caution when soliciting feedback outside of the use context. We repeatedly received comments from parents about the games' visual design not being up to date, being boring compared to Nintendo games, not colourful enough, etc. despite their children clearly enjoying the same games and wanting to play them again and again. While a simple and modern visual design, tested in a rather empty room might seem boring, it might stand out within a visually crowded environment, attracting attention through movement, where a more intricate design might be 'too much of the same'. Here, again, adults tended to compare the installations with common PC-based or video games, and to focus on the graphics as what they believe attracts children.

3.5 Trade-offs and Experiences with our Study Setup

Our motivation for running the majority of the user studies in the Kelvingrove museum library was to create a setup that is less intimidating than having to visit researchers in a University, and to have participants experience the installations in a museum-related setting. The question is how far this was successful and whether the effort involved was worthwhile. While our experiences were not wholly negative, arguably the hurdles we had to surmount in conducting the studies in this way slightly outweigh the benefits gained from using this set-up. Having said this, our approach to simulating the museum group situation by inviting young families was very successful.

To some extent we believe that the setting helped to make participants feel at ease, although this is impossible to verify. This was more important for the elderly adult user group, whereas children quickly ignored the researchers regardless of setting once they started to play, and parents were simply happy to see their children entertained. Testing fun games was certainly an advantage in this respect and helped recruiting participants. Running the studies in a museum meant that parents could combine our study with a museum visit. More importantly, it enabled us to ad-hoc recruit some participants. Disadvantages of working in the museum library involved disturbances by other visitors and overall noise levels. Moreover, security was an issue, since we could not lock the door behind us, and had to carry equipment across town. In addition, being restricted to museum opening times in combination with being reliant on school-attending participants meant that we often had to run studies over the weekend. In contrast, sessions with the table-based game could take place in the late afternoon and early evenings. Some of our reflection furthermore revealed that due to the by-invitation status of sessions, participants might interpret the installations as stand-alone applications, it being too hard to imagine these being part of a larger museum visit experience.

In the ideal case, one would test installations in the museum. Yet this was still being rebuilt and refurbished at the time. Alternatively, one might install prototypes in a children's activity room within a museum, and enlist participants by having the sessions announced as children's activities. This would at least partially enable observation of how strangers interact with each other, and also create a relaxed atmosphere. But this requires a much larger group of researchers for running studies. We were able to conduct our study sessions with only one or two researchers being present, going through consent forms, observing game play and conducting interviews sequentially. In a game-lab setup (Isbister and Schaffer 2008), several researchers would be needed to conduct and supervise all these activities going on simultaneously.

While this is somewhat disillusioning, it was only by taking the effort that we could come to assess these tradeoffs. The aspect of our study that proved to be most important for its success was the focus on family and group interaction and our decision to invite larger family groups, in combination with the effort taken to recreate core aspects of the installation technology (screen sizes, touch-ability, the circular multi-button table). While there are questions left open as to whether this provides us with insights into all of the

potential issues relevant in social situations in the museum, it was essential in order to determine if e.g. the multiplayer games on the table would be fun to play in different group sizes or whether more players would result in confusion, and whether siblings could participate and collaborate in the touchscreen games.

4 Outlook and Conclusion

We have discussed some of our experiences with trying to emulate 'natural' family interactions in a formative user study of prototypes for museum installations. We have focused on the questions that our study left open, in regard to whether groups will behave similarly in the museum, and whether our setup enables us to predict additional interaction issues that result from the ecological setting in the museum situation. In our future research, we conduct observations in the new RBB museum, and will compare these with the evaluation trials.

Acknowledgements

We thank Mary Hudson/Stone from the National Trust for Scotland, the Kelvingrove Art Gallery and Museum in Glasgow, all of our participants, and to Patrick Tobias Fischer and Andrew Smillie for helping to build the table.

References

- Allen, S. (2004), "Designs for Learning: Studying Science Museum Exhibits That Do More Than Entertain", *Science Education* Vol 88(S1), S17-S33.
- Bell, G. (2002) Making Sense of Museums: The Museum as 'Cultural Ecology'. Intel Labs.
- Falk, John H., Lynn D. Dierking and Marianna Adams. "Living in a Learning Society: Museums and Free-choice Learning." A Companion to Museum Studies. Macdonald, Sharon (ed). Blackwell Publishing, 2006. Blackwell Reference Online. 05 July 2008
- Gutwin, C. and S. Greenberg (1998), "Design for Individuals, Design for Groups. Tradeoffs between Power and Workspace Awareness". In Proceedings of CSCW'98, 207-216. N.Y.: ACM.
- Heath, C., and vom Lehn, D. (2002), "Misconstruing interactivity". *Proc. of Interactive Learning in Museums of Art*. V&A Museum
- Hornecker, E. and Stifter, M. (2006), "Learning from Interactive Museum Installations About Interaction Design for Public Settings". *Proceedings of OzCHI 2006*. N.Y.:ACM, 135-142
- Hornecker, E. (2008), "'I don't understand it either, but it is cool" Visitor Interactions with a Multi-Touch Table in a Museum". *Proc. of IEEE Tabletop 2008*. 121-128
- Hornecker, E (2010), Interactions Around a Contextually Embedded System. *Proc. of TEI 2010*. ACM Press 169-176
- Isbister, K., and Schaffer, N. (2008), *Game usability: advancing the player experience*. Morgan Kaufmann

Isbister, K. (2010) "Enabling Social Play. A Framework for Design and Evaluation". In: Bernhaupt, R. (ed) *Evaluating User Experience in Games. Concepts and Methods*. Springer

Kelly, L., Savage, G., Griffin, J., and Tonkin, S. (2004) *Knowledge Quest: Australian Families Visit Museums*. Australian Museum and National Museum of Australia

Macdonald, Sharon (Ed). *A Companion to Museum Studies*. Blackwell Publishing, 2006.
Blackwell Reference Online. 05 July 2008

Marshall, P., Morris, R., Roger, Y., Kreitmayer, S. and Davies, M. (2011), "Re-thinking 'multi-user': an in-the-wild study of how groups approach a walk-up-and-use tabletop interface". In *Proceedings of ACM CHI'11*, ACM Press, to appear.

Monahan, P. Letting Visitors Surprise Us. Proc. Of International Workshop: Re-Thinking Technology in Museums: Towards a New Understanding of People's Experience in Museums 2005, Limerick

Pemberton L. and R. N. Griffiths. (2003), "Usability Evaluation Techniques for Interactive Television". in Stephanidis C. (Ed.) *Proc. of HCI International 2003. Universal Access in HCI, Volume 4*, Lawrence Erlbaum Associates, 882-6.

Sanford, C., Knutson, K. and Crowley, K.(2007), "'We Always Spend Time Together on Sundays': How Grandparents and Their Grandchildren Think About and Use Informal Learning Spaces". *Visitor Studies* 10:2, 136-151

vom Lehn, D., Hindmarsh, J., Luff, P. and Heath, C. (2007), "Engaging Constable: Revealing art with new technology. In *Proc. of CHI'07*. ACM. 1485-1494