PROJECT DELIVERABLE REPORT

Usability of Music for Social Inclusion of Children

UMSIC

Grant Agreement number:	224561
Project acronym:	UMSIC
Project title:	Usability of Music for Social Inclusion of Children
Funding Scheme:	Collaborative Project
Start date of project:	01 September 2008
Duration:	36 months

Deliverable reference number and title:
D3.5b

Due date of deliverable:
31/05/11

Actual submission date:
31/05/11

Organisation name of lead contractor for this deliverable:
University of Central Lancashire (UCLan)

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Additional information:

Author/s:	Lorna McKnight, Netta Iivari, Janet Read, Diana Xu
Beneficiaries contributing to the deliverable:	UCLan, Oulu
WP contributing to the deliverable:	WP3
Estimation of pm spent on the deliverable:	3.5
Nature of the deliverable: Report, Prototype, Demonstrator, Other	Report
Total number of pages:	22

Have ethical issues been taken into consideration in this deliverable (Yes /No)?	Yes

Please give a short summary (5- 10 lines) how the ethical issues have been considered in this deliverable, when working with children? (In addition, see the ethics check list prepared by Prof. Simon Rogerson available on Optima folder Ethical issues.)

If the answer to the question mentioned above is no, please provide a short explanation why ethical issues have not been considered?

Work conducted for this deliverable has involved a considerable amount of contact with children, conducted at the universities of UCLan and Oulu. All of these have followed the ethical guidelines outlined by the UMSIC group, as well as specific ethical working practices of universities involved.

All contact with children at UCLan is assessed by the UCLan ethics board, and is covered by the ChiCI Group's code of conduct. Children are treated as volunteers in the studies and are able to withdraw at any time, and it is stressed to them that they will not 'fail' a task if they cannot complete it. Where possible, all children in a class are given the chance to try any technology brought into a classroom, even if their data cannot be used, so that no-one is excluded – if this is not possible, care is taken to give them a task of equal or greater 'fun'.

Data kept from the studies is always anonymised – only ages and genders are recorded along with the results, not names, and where photographs or videos are used to record activities care is taken to omit faces or identifying information if they are to be published. Photography and video are only used if the school holds consent from the pupils' parents for this type of activity, and if the children do not object.

Any researcher who has individual or unsupervised contact with the children is subject to an enhanced Criminal Records Bureau (CRB) check, as is the standard working practice in the UK. Records are held of all activities conducted, including the class that participated, the researchers that conducted it, who designed the activities and the research questions behind it.

Key words:
Usability, interface, evaluation, children, JamMo.

Abstract (10-20 lines) summarising the content and results presented in the deliverable:

This report presents a summative evaluation of the two JamMo products as of 31st May 2011. The report focuses on two main areas: usability (ease or efficiency of use), and user experience (enjoyment or engagement in use). The results are drawn from a number of studies conducted with children between the ages of 4 and 10. Overall, over 100 children were involved in the evaluation of JamMo 3-6, and around 60 for JamMo 7-12. A brief description of each study is presented, before a summary is given of the main usability findings relating to JamMo interface components. The evaluation identifies a number of usability issues that remain in the software, but overall user experience was reported as positive for the final evaluations.

Description of the deliverable content:

1. Introduction

Work package 3 is concerned with the design of appropriate technology for the target groups within UMSIC. It has proceeded in three phases: the first was to establish requirements, the second was to carry out studies around usability (e.g. Hanna et al. 1997, Markopoulos et al. 2008) as and when prototype applications were developed, and the third phase, the one described here, is to ensure that the product is fit for purpose for the target, and all other possible, user groups. To that end, the work of this work package is about evaluating the technology, and the technology solutions of their usability, for their attractiveness to children and for their general fit to the users. It is not about evaluating the effectiveness of the JamMo suite of products for inclusion, for language improvement or for music learning – all these evaluations are included in Work Package 9 which is the work package intended to address impact.

Usability (e.g. ISO 13407, 1999) is generally described in terms of effectiveness, efficiency and satisfaction. It was traditionally concerned with the evaluation of work systems for their fitness for purpose and is mainly concerned with the extent to which a technology product is able to assist the user to fulfil a goal based task.

Children, as users of technology, are less inclined to be goal-centred and are often more interested in playing with or exploring products – for this reason, usability evaluations with children tend to take a different focus than those intended for adults.

Usability is generally evaluated either during an iterative product development process (formative evaluation), or at the end of a product build (summative). The former type of evaluation indicates bugs to fix, suggests design makeovers and highlights areas of poor performance; in the latter stages, a usability evaluation is rather more focussed on providing a benchmark score that allows a design team to see to what extent their product works. This benchmark can assist the development team in identifying the limitations of, as well as the possibilities for, the product that has been built.

In UMSIC, there are a range of products; these include the two main JamMo music products that are situated on the Nokia N900, the collaborative products that include paired composition and workshop sessions, the orientation games associated with the software suite for ages 7 – 12, the virtual instruments devised by the team from Systema and the online community where JamMo users can interact and communicate.

The focus in UMSIC is on evaluations of mobile technologies in situations where they may be used collaboratively with a particular enthusiasm for understanding how best to evaluate mobile technologies with children in situations where the technologies might be in the same place, in one place but over a length of time, in different places, and in collaborative use. Additionally, the UMSIC team have sought to develop methods that can be easily adapted for, and used with, children who are either newly immigrant, with limited language skills, or suffering from attentional difficulties as manifest in ADHD. Rather than develop methods that specifically suit these two user groups, the aim has been to devise methods that suit all children and in particular children with these needs.

To meet this requirement, the aim is for every evaluation to be:

- Short lasting (to ensure attention is kept)
- Fun (as a motivation)

- As language free as possible (to allow all to take part)
- Non threatening (so as not to cause anxiety)
- Easy to administer in mobile and classroom contexts (to fit the use scenarios)

The evaluation methods chosen will fit with the evaluation framework specified in D3.4. Thus when needing to evaluate for fun, methods will be chosen which are suitable for evaluating fun, and likewise when evaluating usability there are appropriate methods for this. A categorisation of suitable methods has been begun, leading on from D3.4 and D3.2, and will help inform the methods chosen in D3.5b. In line with the framework, a proposed PLU weighting has been suggested for the JamMo products of [P:60%; L:15%; U:25%], and this will be taken into account when planning the studies, although this will also help to act as a test of the framework.

A literature review of methods revealed several methods that in part fitted these requirements. For satisfaction measuring, these included:

- The Fun Toolkit (short, fun, language free, non threatening, easy to administer) – discovers satisfaction and user experience (Read, 2008)
- Drawing of experience (fun, language free, non threatening, easy to administer) – discovers understanding of the system (Xu et al., 2009)
- Picture cards (fun, language free, non threatening easy to administer) – discovers usability and user experience problems (Barendregt, Bekker & Baauw, 2008)
- This-or-That (short, language free, easy to administer, non threatening, simple to respond) – discovers preference (Zaman, 2009)

To determine the efficiency of the JamMo products, i.e. the extent to which the products do as they are supposed to without error and without frustration, the following methods were considered:

- Gathering log data of use (to include errors and slips) during defined user tasks
- Capturing video data of use (as above)
- Capturing think aloud data of use (where the child talks about what they are doing)
- Gathering data from exceptions

To determine the effectiveness of the JamMo products, it may be important to determine the extent to which the technology allows for inclusion, for music sharing and for conversation. This is a technology-centred, rather than a child-centred (which is in WP9) view of inclusive technology and is an area that crosses between WP3 and WP9. Methods that could be used in this space would include:

- Protocol analysis
- Conversational analysis
- Analysis of turn taking
- Questionnaires or interviews about sense of inclusion during use of JamMo products

A variety of methods were used in this report, depending on the age and ability of children taking part, the resources available in the school or laboratory where the studies were conducted, and the status of the software that was available. For example, JamMo 3-6 was developed to a stable level earlier than JamMo 7-12, meaning that more user participation studies could be conducted with JamMo 3-6, whereas for JamMo 7-12 the software was still in development until a later date, and so fewer user participation studies could be conducted.

Overall, over 100 children were involved in this phase of evaluation of JamMo 3-6, and around 60 for JamMo 7-12. This is excluding studies conducted for the usability report D3.5a, pilot studies,

and studies conducted to inform the use of methods, which also used JamMo 3-6 and 7-12 with children, but which are not discussed here.

This deliverable was originally intended to be a report on the finalised software 'after use in the field', as stated in the UMSIC project proposal. Due to the ongoing development of the software however, this report will instead present a summative usability evaluation of the software as it stands as of 31st May 2011, while additional findings on the usability of JamMo after more longitudinal studies of the software in use will feed into WP9 (exploring the impact of JamMo). Evaluations conducted for this deliverable therefore began as formative evaluations, with usability issues and software errors feeding back to the development team, before the final summative evaluations were conducted. The final evaluations were conducted using the latest version of JamMo available, as agreed with the development team, which was version 0.7.24, and which although still under development was considered to be stable and suitable for use with children.

2. Ethical issues when working with children

Work conducted for this deliverable has involved a considerable amount of contact with children (over 100 children from three schools/pre-school centres), mainly conducted at the universities of UCLan and Oulu. All of these have followed the ethical guidelines outlined by the UMSIC project, as well as the specific ethical working practice of each university involved. There is no sensitive data being gathered and the ethical concern has been to ensure that the children are protected from psychological (potentially caused by activities being too difficult or by frustrations with the technology being unstable) and physical harm (potentially caused by poor management of activities or by the use of unsafe products) and that they are treated with fairness and respect.

Before the evaluations

All contact with children at UCLan has been assessed by the UCLan ethics board, and is covered by the ChiCI Group's code of conduct. Any researcher who has individual or unsupervised contact with the children is subject to an enhanced Criminal Records Bureau (CRB) check, as is the standard working practice in the UK. Records were made of all activities conducted, including the class that participated, the researchers that conducted it, who designed the activities and the research questions behind it.

Before the evaluations the school teachers were contacted, and any children with disabilities were reported. According to the size of the class and if any special attention required, the evaluations were designed to allow enough time and fair treatment to all children. We made sure each session was packed with fun activities by piloting it to a small number of children first.

During the evaluations

Children were treated as volunteers in the studies and made able to withdraw at any time, and it was stressed to them that they will not 'fail' a task if they cannot complete it. Where possible, all children in a class were given the chance to try any technology brought into a classroom, even if their data could not be used, so that no-one should be excluded – if this was not possible, care was taken to give them a task of equal or greater 'fun'.

The evaluations took place in school breakout rooms and classrooms, since children felt more at ease in their familiar environments. Children were normally paired with a working partner (a

system already in place at the schools), which helped some children to overcome shyness and encouraged them to help each other and to be more involved with the activity. They all showed great enthusiasm and engagement while working in groups.

After the evaluations

Data kept from the studies was always anonymised – only ages and genders were recorded along with the results, not names, and where photographs or videos were used to record activities care was taken to omit faces or identifying information if they are to be published. Photography and video were only used if the school held consent from the pupils' parents for this type of activity, and if the children themselves did not object.

3. Description of studies conducted

This section provides a brief summary of studies conducted as part of this deliverable, before the summarised results are presented in section 4.

JamMo 3-6

1. 3-method comparison (July 2010)

Software: JamMo 3-6 version 0.5.1
Hardware: Nokia N900
Participants: 24 UK children, aged 7-10.
Methods used: Smileyometers, This-Or-That, Laddering

This first evaluation study (previously reported in interim document 3.5b) was conducted using 3 methods, to compare between them. These methods were: Smileyometers (a survey method from the Fun Toolkit; Read, 2008), This-or-That (a pairwise comparison rating; Zaman, 2009), and Laddering (a structured interview technique; Zaman & Abeele, 2010).

For this study, 24 children from two schools in Preston, UK took part as part of a day of activities held at UCLan. Of these children, 12 were from Year 3 (age 7-8) and 12 were from Year 5 (age 9-10). The children each played 3 games on a Nokia N900, and completed all 3 evaluation methods for each game. One of these games was JamMo 3-6, and for this study the children focussed only on the composition game, using the city theme. The study lasted approximately 20 minutes for each child.

2. Think aloud & interviews 1&2 (Spring 2010)

Software: JamMo 3-6 version 0.6.8 and higher
Hardware: Nokia N900
Participants: 14 Finnish children, aged 5-6.
Methods used: think aloud, interviews

This study was carried out with children in Oulu. The usability tests involved thinking aloud and interviewing the children before, during and after the test sessions (see e.g. Dumas and Redish, 1993; Rubin, 1994). Heuristic evaluation (Nielsen, 1994) was carried out as well as a pilot test before the actual test sessions, as well as the pilot study involving children. Free testing was carried out with the first 4 children, but during the remainder of the tests task based usability testing was used instead of free exploration because of the instability of the application being tested. Also active intervention was in use. The first 10 tests were individual tests, while the last 4 tests relied on peer

tutoring method, i.e. one child teaching another to use the application (cf. Markopoulos et al., 2008).

3. Observations, Fun Toolkit & interviews (Spring 2010)

Software: JamMo 3-6 version 0.6.8 and higher
Hardware: Nokia N900
Participants: Finnish children aged 5-6
Methods used: observations, Fun Toolkit, interviews

In this study fun evaluations involved the use of a Funometer, Smileyometer, Fun-Sorter and Again-Again table (see Read, 2008). The Funometer and Smileyometer were combined to make it more suitable for children of this age (5 to 6 years). In addition to the use of these instruments, the children were observed during the usability testing as well as during music evaluation sessions that were organized together with music education researchers. Furthermore, the children were interviewed before, during and after the usability tests and they were asked to draw something fun. A careful analysis of all the results gained was carried out, comparing the results gained through the use of these different methods as well as identifying aspects that supported or hindered fun related to the application use.

4. Observations and Smileyometers (March 2011)

Software: JamMo 3-6 version 0.6.17
Hardware: HP Pavilion tx2000 touchscreen laptops
Participants: 36 UK children, aged 8-10
Methods used: observations, Smileyometers

This session formed part of a day of activities at UCLan with two classes from a school in Preston, UK, which coincided with a visit from all the UMSIC partners. In the first half of the day, children took part in a series of activities for approximately 20 minutes each, one of which was a JamMo evaluation. For this, children played the JamMo 3-6 game on a touch-screen laptop, with observations from a researcher. The screens were also video recorded for more detailed evaluation later.

In this activity, children were not directed, but were allowed to freely explore the game. They worked in pairs so that conversation could be captured, and there were 3 pairs around each table at once. After playing, a Smileyometer (see Read, 2008) was used as a simple assessment of fun for the activity.

5. Observations (April 2011)

Software: JamMo 3-6 version 0.6.17
Hardware: Nokia N900
Participants: 9 children, aged 6-7
Methods used: observations

This session formed part of a half-day of activities at UCLan with one class from a school in Preston, UK. Children from the class took part in a series of activities, with 9 children taking part in an observation study using JamMo 3-6 on a Nokia N900. Children were given tasks to complete using the JamMo 3-6 game, while they were each observed by at least two researchers, who made notes during the session or referred to a video recording to make notes after the session. Children worked individually or in pairs.

6. Peer game (April 2011)

Software: JamMo 3-6 version 0.6.17
Hardware: HP Pavilion tx2000 touchscreen laptops
Participants: x UK children, aged 6-7
Methods used: observations & questionnaires

This session formed part of the same half-day of activities as study 5 above, with children who did not take part in the observation study instead taking part in a test of the peer game. Children worked in pairs on a laptop, which was then paired using the peer-to-peer network with another laptop situated on the other side of the room, for children to play the co-operative networked pair game in the JamMo 3-6 composer. This was used as a test of the pair game with developers and usability researchers present, which helped to find bugs or features to improve in the networked game, and also usability and user experience issues with the game, and an understanding of whether and how the children could use it to collaborate.

The children played the collaborative composition game while each pair was observed by a separate researcher. Afterwards, they completed a questionnaire using Smileyometers to rate the game on usability and user experience factors. Notes were also taken by the researchers.

7. Pre-school Observations (May 2011)

Software: JamMo 3-6 version 0.7.23
Hardware: Nokia N900
Participants: 5 UK children aged 3-4
Methods used: observations

This study involved children from a UK preschool centre, who took part in an informal session using JamMo 3-6. This aimed to gather information from the youngest end of the age range of JamMo, but due to the children's age this study was designed to be as simple and non-threatening for the children as possible. Two researchers visited the preschool centre, and the preschool staff selected children for the activity who they felt would be able to take part.

The children were allowed to play the game freely, with no specific instructions, but assistance was given by the researchers when needed. Two children played the game at once, each on their own device, meaning that the researchers could not only observe the actions on the screen but also listen to conversations between the children. Only a small number of children took part, but they were allowed to play with the JamMo game for as long as they wanted, which took up to an hour per child. The researchers made notes on the session afterwards, as note-taking during the session was felt to put pressure on the children.

8. Final evaluation (May 2011)

Software: JamMo 3-6 version 0.7.24
Hardware: Nokia N900
Participants: 20 UK children aged 5-6
Methods used: observations, questionnaires, JamMo software logging

This study aimed to act as the final evaluation of JamMo 3-6. The method was informed by findings from previous studies, using a mixture of observations and questionnaires. The children were seated in a quiet room with no other activities taking place, and used the N900 with finger or stylus as they chose. Before playing the game, the children were only told that it was a music game, and were left to explore and play as they liked, but they were also told that the game contained a teddy bear who would tell them how to play the game, and the mentor's first introduction message was played to them, to ensure that they knew the mentor was available to give help and instruction if desired. The

children each played the game for 15-20 minutes before the questionnaire was completed by the children with assistance (e.g. with writing) from the researchers.

Observations were conducted by trained researchers, aiming for 1 or 2 observers per child, so as to be able to capture as many on-screen actions as possible, and the children worked in pairs (each with their own N900) with a partner who was well known to them, so that they would also talk to each other about the tasks. This allows the observers to gather more information about the children's understanding and experience of the software, in a similar manner to the Peer Tutoring method (see Xu et al., 2007), and also makes the children more comfortable than working individually with a researcher. The questionnaire was in two parts: the first contained screenshots of the game, which the children labelled, to show how much of the interface they understood after playing; the second part contained questions similar to those in the questionnaires being developed for WP9, to allow for comparison with later studies. This used Smileyometers (Read, 2008) to rate fun, ease of use and how well the game worked on the day, as well as asking if they would want to play the game again, and what the best and worst parts of the game were.

9. Additional evaluation of JamMo usability for teachers (April 2011)

Software: JamMo 3-6 version 0.6.17
Hardware: Nokia N900
Participants: 2 primary school teachers
Methods used: think aloud, observations, interviews

As an addition to the evaluations described above, a small investigation was also carried out as to the usability of the JamMo game for teachers. As teachers are considered a secondary user of the software, they would be expected to teach children how to use the game, and support them in their use of it, as well as potentially using it to prepare lesson plans and activities for the children to take part in. Suitability of the game for teachers in music education will be addressed in WP9; this study looked simply at the ease of use and understanding of the software for teachers, as well as their impressions of the software and how easily they could imagine using it with children.

Two primary school teachers (categorised as 'late adopters' of technology) were invited to take part in this study. Each worked individually on separate days. For the main task, the participants were asked to use particular features of JamMo 3-6, using the 'think aloud' method for protocol analysis. A researcher also observed the participants interacting with the software. Pre- and post-test interviews were also conducted to gather additional qualitative information.

JamMo 7-12

1. User test and questionnaires (March 2011)

Software: JamMo 7-12 version 0.6.17
Hardware: Nokia N900 and HP Pavilion tx2000 touchscreen laptops
Participants: 36 UK children aged 8-10
Methods used: observations and questionnaires

This session formed part of a day of activities at UCLan with two classes from a school in Preston, UK, which coincided with a visit from all the UMSIC partners. In the second half of the day from study 4 using the 3-6 game, children took part in a 45 minute activity using JamMo 7-12. This was the first large-scale test of JamMo 7-12, being used by a whole class at once, and acted as a capacity test of the system as well as allowing all partners to see JamMo in use by children.

For this study, children were allowed to play freely with JamMo 7-12 on the Nokia N900s, working in pairs. One touch-screen laptop per table of 4 children was also then made available, due

to instabilities with the N900s. Researchers circulated around the room making observations and assisting the children where needed. After the children had finished playing, researchers administered a questionnaire, which used a rating scale to measure usability and user experience measure, and which acted as a pilot for a questionnaire being developed for later studies in WP9. A debrief session for the researchers afterwards allowed notes and observations to be shared.

2. Usability tests and Interviews (April 2011)

Software: JamMo 7-12 version 0.6.14-2
Hardware: Nokia N900
Participants: 3 Finnish children aged 13-15
Methods used: usability tests and interviews

Usability tests with 13 to 15 years old kids have been carried out involving 3 kids, two girls and one boy. Before the tests, a cognitive walkthrough was carried out (Wharton et al., 1994), resulting in the identification of some potential usability problems. After the cognitive walkthrough evaluation, task based usability testing was carried out, involving interviews before and after the test sessions. User experience was inquired with the use of a form with emoticons. The focus was on usability, but also on user experience more broadly as well as on fun. The material was carefully analyzed after the tests and the usability problems were identified.

3. Expert evaluations (April 2011)

Software: JamMo 7-12 version 0.6.17
Hardware: HP Pavilion tx2000 touchscreen laptops
Participants: 3 expert evaluators
Methods used: heuristic evaluation

As a part of the studies being conducted, an expert evaluation was also conducted on JamMo 7-12. This used the heuristic evaluation inspection method, using experts in usability, and used the heuristic set developed for use with computer games by Pinelle et al. (2008). First, the evaluators were all briefed with the method, and made familiar with the heuristic set being used, and examples of each heuristic in use. The evaluators each used the JamMo 7-12 game, and noted problems that contravened the heuristics. Then, following the method, all evaluators met to discuss the problems they had discovered, in order to decide on a combined set of usability problems encountered, along with severity ratings for each problem.

4. Observations & questionnaires (May 2011)

Software: JamMo 7-12 version 0.7.24
Hardware: Nokia N900s
Participants: 21 UK children aged 9-10
Methods used: observations, questionnaires, JamMo software logging

This study aimed to act as the final evaluation of JamMo 7-12. The method was informed by findings from previous studies, using a mixture of observations and questionnaires. Before playing the game, the game was demonstrated to all the children on a laptop to explain how to use it, due to the absence of an English-speaking mentor in the software at this time. The participants then each played the JamMo game individually on the Nokia N900. The children were seated in a quiet room with no other loud activities taking place, and used the N900 with finger or stylus as they chose. Three to six children took part in the activity at one time. When more than three children took part at once, the children were seated in pairs to allow for conversation, and to make the children more

comfortable; the children were each using headphones to improve system performance so conversation was limited, but did occur.

Each child was observed by a researcher who took notes throughout the session. The researcher could also give help or assistance if asked for by the children, but otherwise aimed not to interfere. After the children had finished playing, they each unplugged their headphones and played back their finished composition, so that it could be heard by the researchers and the other children. They each then completed a questionnaire, with assistance from the researchers. This used a similar format and questions to those used in study 8 for JamMo 3-6, and to the questionnaire developed for WP9.

4. Summary of usability and user experience findings

The data from all the sections described above were analysed, gathered together, and grouped into themes. The combined findings for each individual component of JamMo are therefore described separately below.

JamMo 3-6

Door screen

The 'door' screen acts as the opening screen of the game. This was informed by children's designs, and aimed to act as a welcoming portal for children, and to increase a sense of mystery and adventure.
Studies in the UK had previously shown that children and expert evaluators were unsure about the purpose of the door (see D3.5a), but studies in Finland indicated that the door used for entering the application was considered to be a good choice.
This screen is only encountered once briefly in children's interaction with the game, so usability issues related to this screen are considered to be of minor importance.

Login screen:

The gesture-based login screen was designed to provide access to children's individual profiles, through use of a visual password using connections between dots.
Early studies of this in Finland suggested that login was unclear for the children. However, this may be due to the fact that this feature is still in development, and so because of this, this function was removed from the version of JamMo that was used for the final evaluations, and so will not be addressed here. This feature will however be evaluated separately for WP7 and WP9.

Mentor

The 'mentor' in JamMo 3-6 currently takes the form of a teddy bear, who gives instructions, feedback and encouragement for playing the game. This was considered an essential feature, particularly to support children with attentional difficulties.

In usability terms, the placement of the mentor was found to be consistent, and children mainly understood what the purpose of the bear was. They could hear his speech and were noted to be listening to him and responding to what he said – often they used phrases that the bear had just said, and when the bear said that the composition was saved in the cupboard one child immediately responded "cool, thanks". However, some children ignored his speech, or still had to ask for instructions after listening to him. Not all children understood how to activate or disable the bear though, and often encountered errors when clicking the character to make him go away, by repeated clicking, which made the character re-appear unintentionally as a second click was interpreted by the system as re-activating the character.

In user experience terms, some children found the mentor encouraging, and enjoyed his feedback. When it was explained to children that the game contained a teddy bear that would tell them how to play it, several children seemed excited and happy by this idea. However, in use, a lot of children found the mentor highly frustrating. Most of this was due to difficulties with disabling the mentor, as described before, but also with the fact that he re-appeared again too often, interrupting their work as they were discussing the task with a friend or deliberating over which loop to choose. Children would often be heard to say "shut up!" and "move away!" when the bear re-appeared, and one child was lamenting for some time "this is the stupidest bear ever because he won't even work... work, bear, work! Could you work please? Come on!" With the pre-school children in the UK in particular, they developed a game of 'kill the bear', in which they hit the screen repeatedly as hard and fast as they could, to try and make him go away. The feedback he gave was also felt to be excessive, and in some cases meaningless, for example when he says "very good, you sang very well" after the singing game, even if the child never sang.

However, it is important to note that this issue has already been addressed by the development team – while the inclusion of a mentor was felt to be important, it was always the intention that it should be able to be disabled if needed, and a mechanism for this has been included in later versions of JamMo. The main outstanding usability issues remaining with this feature therefore are ensuring that the mentor's comments are as useful and helpful as possible, and ensuring that the mentor does not re-appear if accidentally clicked twice to disable.

Home screen

The home screen is the main menu for the JamMo 3-6 game. Derived from children's drawings, this represents a house (continuing the metaphor of the 'home' screen), and contains the menu items for the singing game, the composition game, the cupboard, the difficulty settings, and the logout button.

In general, the main screen was found to be clear and understandable, with the options clearly visible to users. Findings for the individual components on this screen will be addressed individually.

Game menu icons

These buttons are navigational icons, leading to the singing game and the composition games respectively, which are described separately later in the document.

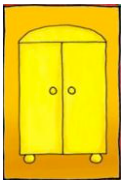
Across several studies the children were able to identify the pictures used for the singing game and composition game menus, for example 'boy singing'. However, quite often across studies the children could not connect the picture with the activity it represented. They often referred to the options as 'the girl' or 'the boy', rather than songs or compositions, saying "let's go on the girl" or "go to the boy". They would ask for help, saying "I want to get onto the places", showing that they could not connect the image with the screen it leads to. When asked the meaning of the icons after playing, they could usually remember which activities were behind each button (for example saying 'if you click the boy you go to a place that plays music, but if you click the girl you go to a place and make music'), but did not make comments suggesting that they understood the connection between the icon and the activity.

Difficulty settings icon

This button allows users to choose between 'easy' and 'advanced' settings on a following screen, and also acts as a status icon to indicate which setting is currently in use, between a single note representing 'easy', and a series of notes (called a 'multi-note' by the mentor) to represent 'difficult'.

Children could change the settings if this was explained to them. However, most did not recognise it as a button to choose options, instead seeing it as just decoration, and the result of changing the settings was not obvious to the users. In the short test with teachers, they did not understand this button or what the image represented (other than 'music').

Cupboard icon



This icon leads to the cupboard screen, as described separately later in the document.

While children recognised this as an item of furniture, they did not seem to understand its place in a music-making game. One child in the final evaluations asked "why is there a wardrobe?", and several other children referred to it as a 'wardrobe' or 'drawer type thing', suggesting that the image resembles a wardrobe (for clothes) rather than a cupboard (for storage).

Logout icon

This navigational icon represents 'logout' or 'quit', illustrated by the mentor walking away through a door, and leads to a confirmation screen where the user can either confirm or cancel the action.
Many of the children identified the purpose of this icon correctly. However, there were also several cases of users accidentally quitting the game through choosing the wrong option at the confirmation screen, wanting to explore all interface components, or through pressing the icon twice. This icon is therefore perhaps not clearly recognisable as quitting the game, although this is confirmed by the mentor's speech if the mentor is activated and the children are attending to his speech.

Singing game

The 'singing game' is one of the two main activities in JamMo 3-6, involving a karaoke-style game where children are encouraged to sing along with a recorded song of their choice, and then listen back to a recording of their own singing. The game starts at the song options screen, and continues onto the song playback screen, each discussed below.

Song options screen

On this screen, the user can select from a number of icons representing songs, each of which leads to a different song playback screen. The user can also change the language settings and choose whether they are playing individually or with an adult.

Song icons

Several studies across the UK and Finland suggested that the song icons were not easy for the children to recognise, were not inspiring for the children and they could not identify the songs from the image. Some of the images were not thought to closely represent the songs they led to, such as the sun image for the song 'Lightly Row', and the raindrops for 'Fox You Stole My Goosey Gander'. On at least one occasion a song was wrongly identified, for example a child saw the image of a sheep and said it was 'baa baa black sheep' (another nursery rhyme about a sheep), and then was disappointed when it was not the song they had wanted. They would identify the songs to each other in terms of their picture (e.g. saying "go on the bed") rather than by the song it represented. The children would often explore the content of each of these images, which supported their curiosity, but they would also become frustrated if the image led to a song they did not like.

Game settings icon

A study in Finland suggested that the use of the adult and child image was clear and inspiring for the users. Children in the UK however mainly did not use this option.

Language settings

A study in Finland suggested that the flag icon was clear and inspiring for users. In studies in the UK, the children recognised the image and understood the meaning of the Union Jack flag as 'England', but assumptions were made by both children and teachers that this setting changed the songs that were available (e.g. English songs with the UK flag, French songs with the French flag, etc.) rather than changing the language of the songs that were already visible.

Song playback screen

This screen displays a picture representing the song chosen, while the recorded song plays to the user. The user is expected to sing along (directed by the mentor), and their voice is recorded. The user can then listen back to their own recorded singing, with the same screen displayed.

The children in Finland seemed much more enthusiastic about this game, with studies reporting that the songs were inspiring to the children, and that they liked singing familiar songs, and listening to the songs being sung to them. It was also reported that they enjoyed listening own voices after singing the songs, and that this was an enjoyable activity for them. Some of the children in the UK enjoyed the singing game greatly, singing along with many of the songs, and laughing to hear their voices or their friends' voices playing back (for example showing recognition of the recording by saying to a friend "that's your voice that is!"). However, much more common was for the children to simply sit and listen to the song, without singing along,

even if the song was known to them. Some of this may have been due to the setting of the activity, and the fact that the children did not feel comfortable singing on their own, and this would perhaps not be an issue if situated in a music class or music activity. Many of the children did not realise that the system was recording them, which led to unexpected recordings being made, and which can potentially have serious implications for security and privacy, so this needs to be made much more obvious to the user.

Some of the songs were felt to be too long for the children, both in the UK and in Finland, and they often got bored, and would quit before the song was finished, or be heard saying "hurry up".

The logic and the structure of the game were felt to be simple and understandable for the children, and it was easy for them to learn how to use. However, it was felt that perhaps there were not enough challenges or options available to the children in this game, and they might easily get bored with it. Often children displayed signs of impatience while listening to the songs, especially if they were not singing, and would tap over the screen background or try to click and drag the song cursor, or would exit the game before the song was finished.

Not all of the songs were popular with the UK children across the 3-6 age-group, as well, as the songs used are considered 'nursery rhymes' or pre-school songs in the UK, where children start school at 5. However, children even into the 8-10 age-group were found to like some of the songs, and would listen to the songs, sing along, or sing along with alternative versions with humorous lyrics. Across both the UK and Finland children would get bored or frustrated if they did not know the words of the song, or if the words used were different from the version they knew – for example with 'Old McDonald Had A Farm', this is a very well-known song in the UK, but the order of the verses often changes each time it is sung, and children in one study began to only sing along with the 'E-I-E-I-O' chorus.

At this point in the development there were also several reported issues with the sound quality or volume, causing problems for the user experience of the game. In particular, the children in both Finland and the UK could often not hear their own voices when listening to their own songs afterwards, which has quite a negative effect on the game and the purpose of it.

Composition game

The 'composition game' is the second main activity in JamMo 3-6, and allows the user to select a 'theme' to make music with, which leads them to a different scene containing different types of music. The user can then build their own composition by adding loops of music to a simple sequencer which plays with a backing track.

Theme selection screen

On this screen, the user can choose between the three themes, or scenes – the city, forest or castle scenes. Each of these will lead to a different composition screen.
No particular issues relating to this screen were recorded – children seemed to understand the sub-goal of choosing a 'place to make music' by choosing the picture of the place they were going to. The navigation options on this screen therefore seem clear and understandable for the users.

Composition screen

This screen represents the loop-based composition part of the game, and is different for each of the three themes. Additionally, each theme has three different selections of loops that are available on loading the scene. The loops can be played back by clicking on them, or dragged to the loop track to add to the composition.

Many users could be seen playing the game in one of two non-expected ways. First, rather than building a composition, many users instead chose to play back the backing track, and then 'jam along' by pressing the loop icons during playback. Some users did not understand that they could use the track to build a composition, but others did and still preferred this method instead. Second, many users were seen playing a game of 'filling the track', by adding a number of loops to it, without having listened to them or the backing track beforehand (since the loops do not play back if they are dragged), so they were not engaging in building a composition, although they did often then play back the resulting track, so this was more building through serendipity rather than musical skill. However, most users did understand the main purpose of the game, and generally this design of game was felt to be usable by the children, with a suitable level of challenge for this age-group.

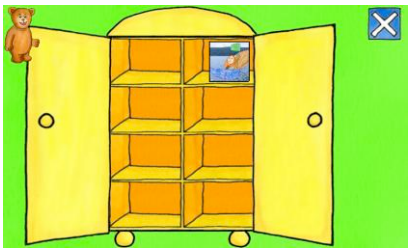
The main usability issue with this game lies with difficulties with the drag-and-drop interaction, which occurred to some extent with a vast majority of all users. Although nearly all of the users understood the interaction method, and understood what they had to do, they had difficulty performing the action. Partly this was due to the children's difficulty in performing the action accurately enough, for example not placing the loop entirely in the loop track, or lifting the finger or stylus off the screen before completing the action. The system lag was another noticeable and frequently occurring source of difficulties though, as the loop did not always follow the user's control closely. Drag-and-drop actions were observed to be easier with the stylus than a finger on the touch-screen, as the children's fingers obstructed the view of the target.

In terms of usability of interface components, the children mainly understood the purpose of the play and stop icons, and only a few had to be told how to play back their composition. The loop icons were mainly understood, but children often tried to drag other items to the loop track as well, such as the castle (part of the background), and the mentor. When aiming to play back the loops, the children did often accidentally drag them slightly, which meant they were not played. The loop track was understood, but the children had difficulty scrolling it, and occasionally removed loops from the track accidentally while attempting to scroll the track. The 'X' icon to close the game was also well understood by users in general, and children recognised it as 'close' or 'exit' – however there were usability problems with this icon, as children did not understand that it would also lead to a 'save' screen, rather than closing the game (indeed some thought it would quit the whole application), and it was not understood that in this context the X also meant 'forwards' rather than 'back'. The saving function as a whole was not well understood by the children, and was the teachers' least favourite part of the game. Also, if this button was unresponsive or the child was impatient, they would press this icon several times repeatedly, and on some cases this led to the child accidentally skipping through all exit buttons in the game and quitting the application unexpectedly.

In terms of user experience, the children greatly enjoyed the music-making aspect and aim of the game, and the musical loops and sounds in the game, all of which they mostly rated highly on Smileyometers, and were noted to be responding to positively (for example saying "cool!" or "listen to this!"). They seemed to enjoy the exploratory aspect of the game, and appeared engaged in the

task of discovering the different loop sounds, often comparing them with their friends and recommending the 'best' ones, and laughing when sounds were played back. The teachers also felt that the music creation was the best feature in the game, and that their children would enjoy it. The pictures and visual elements of the game were also seen to be engaging, with children showing excitement by particular features (e.g. "the robots are cool!" from the older children or "ooh a spider!" from the pre-schoolers). On the whole, the composition game was generally felt to be more engaging and fun for the users than the singing game.

Cupboard screen



The 'cupboard' represents the file storage area of the JamMo 3-6 game, where children's saved compositions and recorded songs can be found. Each saved file is represented by an icon, which in the case of the songs represents the song sung, and in the case of the compositions is chosen by the user from a selection of the loops they have used.

While many children discovered the cupboard (also often called a 'wardrobe' or 'drawer') and used it to discover previously saved items, not many used it to play back compositions or songs of their own. The children had difficulty in recognising their own compositions in the cupboard, and were unable to identify their own creations as several other files had the same icon. There appeared to be a lacking of mapping and consistency: e.g. the singing item last saved was not always on the top of the list in the cupboard. Only a maximum of 8 songs can be saved at a time in their listed order in the *song options screen* with no repetition, and a new singing automatically overwrites the old, which may not be obvious to children. Due to a delay between selecting a file and the file loading, it was difficult for the children to tell which file they had selected if they made many quick clicks on the screen.

Browsing the cupboard was challenging; many users did not realise that the shelf could be scrolled, and even once understood it was difficult to accomplish. There was no indication where the user was in the cupboard (e.g. scroll bars). Files were often selected by accident while browsing. Also, the doors of the cupboard seemed responsive to clicks, so children attempted to close the doors, but this does not have any effect.

In general, the cupboard was not greatly used in the majority of the evaluations, although this may be due to the short nature of the studies, and it may be that it was not necessary in such a situation. However, the teachers felt that the cupboard, along with the rest of the saving mechanism, was the weakest feature of the JamMo game.

General issues with JamMo 3-6

There were several issues which affected the whole of JamMo 3-6 in terms of usability and user experience.

The main problems experienced for all users were system crashes, bugs, or the software not functioning as intended. Due to the software being still in development, some system instability was unavoidable. In the main, most of these problems that were discovered in evaluations were fed back to the developers, and fixed in following versions, or hardware solutions were discovered, such as disabling the wireless capabilities of the N900s, which was found to use processing power and slow the whole application down. System lag remained the largest problem for users, and any delay

between user action and system response caused frustration, and repeated pressing of icons such as the mentor, close icon and loop icons. This did affect user experience, although usually not to any serious extent, and the users were very patient and understanding that they were using developing software. Even children who gave feedback that the game had not worked very well for them said that they would want to play it again, and rated it as highly fun. In the final evaluations, all children said they would want to play the game again, and rated it as highly fun and easy to play. However, it should be noted that children of this age-group are known to respond positively in evaluations, so this finding should be treated with caution, and while it appears an encouraging finding at this stage in terms of usability it remains to be seen how the user experience changes with long-term use. This question will begin to be addressed in WP9 when exploring the impact of JamMo after more long-term use in its intended context.

JamMo 7-12

Main menu

The main menu screen of JamMo 7-12 gives three options: the main sequencer, the orientation games, and the community features.

As the JamMo 7-12 software was still in development for many of the evaluation studies and the English-speaking mentor was not yet available, the main menu was not used much with users in the UK, who instead started the tests from the main sequencer. Studies in Finland however found that the response time after pressing the menu buttons was too long for users. The expert evaluators felt that there was no help features or information on the main screen, that the meaning of the menu items was not clear, and that the mentor should give feedback on being selected – however this may be addressed once the mentor is active in all languages.

Mentor

The English-speaking mentor was not yet available at the time of this report, and so this feature was not evaluated in the UK. However, a few children commented on the character, either asking what it was, or one child said they 'did not like the fluff ball'. With the mentor not being active, the expert evaluators could not see the purpose of the icon.

Orientation games

The orientation games are a feature of JamMo 7-12 designed to support learning of the JamMo features through structured activities, to allow for scaffolded learning. Users are guided through JamMo activities by the mentor, who gives instructions and encouragement, and feedback when actions are performed.

Although this feature had been developed, without the presence of the mentor in English it was felt that the orientation games could not fairly be evaluated in the UK with children, as the tasks would not be understandable. Therefore, this feature has currently only been addressed in Finland.

The users in Finland used the orientation games, but had issues with the wording of the mentor's instructions. The users did not always understand the instructions, for example when the mentor said "listen to the piece and drag melody loops to the track" the users got mixed-up and did not

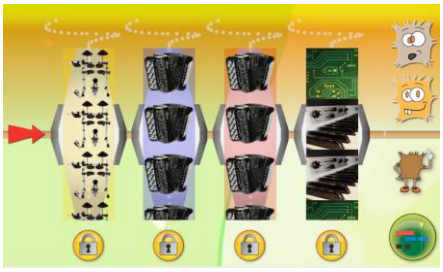
know what to do until they started to drag the melody loops. When ragging a loop and dropping it on the rhythm loop, the mentor's speech "take the loop all the way to the track" irritated the users, one of whom responded with "shut up you creep". Also, when the mentor said "great! Your first composition is ready. It has rhythm and melody", the mentor did not then guide the user to listen to the first composition, which should be performed to enable understanding of the result of their actions. Therefore there appear to be problems with the mentor's speech, but further studies are needed to determine whether the problem lies with the instructions being given or the tasks that are being asked of the users.

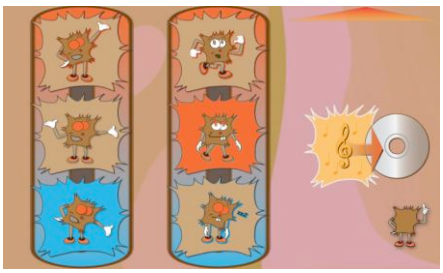
Community features

This area of JamMo 7-12 was designed to hold features such as the user's profile, a discussion forum, sharing of compositions and loops between users, and the users' own file storage area. Although many of these features have been prototyped or developed, they were not felt to be finalised enough to be fairly evaluated at this stage by users.

Sequencer

The sequencer is the main focus of JamMo 7-12, containing the 4-track sequencer, 4 loop wheels and a backing track, as well as options to change the pitch and speed of the track, save in audio or project file format, and play back the final track. The options are split across three scrollable screens: the loop wheels, the sequencer, and the speed and pitch options.





All users seemed to understand the purpose of the game, and used it to create compositions. However, there were issues that were identified through both user participation studies and expert evaluations.

The main usability issue occurring for the JamMo sequencer related to drag-and-drop interaction, particularly of the loop icons. Nearly all users encountered difficulties dragging loops from the loop wheels to the sequencer, which was more difficult than in JamMo 3-6 as the loops had to be dragged from one screen to another; most overcame these difficulties after some time, after the action had been explained and demonstrated to them, but some users struggled with this throughout, and had to ask for help more than once. As with JamMo 3-6, part of the problem was caused by children struggling with the action itself, and releasing the stylus from the screen too soon, but it was also often caused or exacerbated by system lag, and the loops' movements not aligning with the user's actions. A jerking movement as the icons were dragged made it very hard to place them accurately on the loop track. Removing loops from the track was also difficult, without accidentally only moving them.

The users also had trouble understanding the reason for the four different tracks, and why certain loops could not be placed on certain tracks. In the later version of the software used for the final evaluations the four tracks were colour-coded to assist with this, and this seemed to help, but the coloured backgrounds were a bit too subtle to be obvious to users, particularly as the four loop

wheels are vertical, while the tracks are horizontal, which means that the match between the two is not clear.

Manipulating the loop wheels also caused difficulty. Although the children enjoyed exploring the loops and discovering new sounds, it was not obvious to all users or evaluators that the wheels scrolled, and children often seemed to use only the loops that were initially visible in the wheels until they discovered this feature. Even once it was understood, the wheels were difficult to control, either spinning too far, or selecting a loop accidentally when attempting to spin the wheel. The use of the 'crazy' and 'happy' faces to automatically select a combination of loops was not understood by the children, and they often attempted to treat these icons as loops and drag them to the track.

Scrolling the sequencer track also proved difficult, causing accidental removal of loops or pressing options such as the backing track, and the track did not scroll automatically on playback. It was also felt that it was not obvious which part of the overall track was visible in any current view.

The speed and pitch change screen was not well understood by the users or evaluators, and for some users the difference made by the changes was not obvious to them. However some users did use this feature to change the pitch or speed of their compositions.

The option to change the backing track was popular with the children, and they enjoyed exploring the different songs available, with no child choosing to mute the backing track. They had some difficulty in manipulating the scroll buttons on this screen though, and did not all understand the difference between the 'ok' and 'x' buttons.

The metronome feature was not used by children, although they understood its purpose when it was explained. Some children attempted to drag it to the track as though it was a loop.

The main issues encountered in JamMo 7-12 that affected user experience were system crashes and instability. This often made loop dragging more difficult, and occasionally quit the game unexpectedly, losing a user's compositions. In particular in the earlier evaluations with children, the feedback was very negative, mostly relating to the speed of the software and saying that it 'didn't work', and the worst part of the game was that it broke down. In the final evaluations though, many of these issues had been addressed. While there were still occasional issues, and the worst part of the game was often reported to be system crashes, the worst part for many other children was instead the dragging of loop icons. In the final evaluations, JamMo 7-12 was only rated mediocre on ease of use and how well it worked, but was still rated highly on fun, and all users said they would want to play the game again.

Throughout all the evaluations with children however, the music and the loops in JamMo have rated very highly with the users. They nearly always responded that the best part of the game was the sounds in it, or the act of making their own music. This does suggest that this format of game is likely to be popular and engaging with this age-group, and while there are usability problems with the interaction method, the challenges of the game design are appropriate for a good level of user experience. However, as with JamMo 3-6, the user experience after extended use has yet to be determined, and this will be addressed in later work-packages.

5. Conclusion

This report has identified some usability issues with both JamMo 3-6 and JamMo 7-12 after extensive testing with child users, showing the importance of user participation to inform and test usability and user experience of children's interactive products. The main source of usability and user experience problems encountered during testing usually related to system errors and software lag, which are continually being addressed in the latest JamMo releases, and many of which have already been solved, so it is the remaining underlying usability issues relating to the design that are the most use to consider. In particular, the level and format of instructions given by the mentor seem to be an issue to address, although some of this has already been addressed by the development team. The other main outstanding issue seems to be the drag-and-drop interaction used in the game, which caused considerable difficulty, but there was evidence that children began to overcome the difficulties in a reasonably short space of time when the system lag and novelty of the system were no longer an issue, suggesting that this is an issue that can be fixed in later versions of the software rather than a problem with the underlying interaction method.

In general, usability of the two games was good, with children understanding the aim of the games and the options that were available to them, and despite a few problems could use the games to build compositions. User experience however was very good, particularly with later versions of the software where system errors and hardware issues were reduced, and children generally reported that use of JamMo was a positive experience for them, and they greatly enjoyed the music and the sounds, and were enthusiastic about using the device to make music. While this report has suggested improvements that could ideally be made if more time is made available for development, it has also helped to inform areas to be addressed in more detail in the remaining work-packages looking at the impact of JamMo in the classroom.

6. References

Barendregt, W., Bekker, M.M. & Baauw, E. (2008). Development and evaluation of the problem identification picture cards method. *Cognition Technology & Work* 10, 2 (March 2008), pp95-105.

Dumas, J. and Redish, J. (1993). *A practical guide to usability testing.* Norwood: Ablex Publishing Corporation.

Hanna, L., Risden, K. and Alexander, K (1997). Guidelines for Usability testing with Children. Interactions magazine. Volume 4, issue 5, September/October, pp- 9-14.

ISO/IEC (1999). 13407 Human-Centred Design Processes for Interactive Systems.

Markopoulos, P., Read, J. MacFarlane, S. and Höysniemi, J. (2008). *Evaluating Interactive Products for and with Children.* San Francisco: Morgan Kaufmann.

Nielsen, J. (1994). Heuristic Evaluation. In *Usability Inspection Methods,* J. Nielsen and R. L. Mack. New York, John Wiley & Sons, Inc.

Pinelle, D., Wong, N. & Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (CHI '08). ACM, New York, NY, USA, 1453-1462. DOI=10.1145/1357054.1357282 http://doi.acm.org/10.1145/1357054.1357282

Read, J.C. (2008). *Validating the Fun Toolkit: an instrument for measuring children's opinion of technology.* Cognition, Technology and Work, 10(2): p. 119-128.

Rubin, J. (1994). *Handbook of Usability Testing - How to plan, design and conduct effective tests*. New York, John Wiley & Sons.

Wharton, C., Rieman, J., Lewis, C., and Polson, P. (1994): The cognitive walkthrough method: A practitioner's guide. In Nielsen, J., & Mack, R. L. (Eds.), Usability inspection methods, 105-140. New York, NY: John Wiley & Sons.

Xu, D., Read, J.C., Mazzone, E., MacFarlane, S. & Brown, M. (2007). Evaluation of tangible user interfaces (TUIs) for and with children: methods and challenges. In *Proceedings of the 12th international conference on Human-computer interaction: interaction platforms and techniques* (HCI'07), Julie A. Jacko (Ed.). Springer-Verlag, Berlin, Heidelberg, 1008-1017.

Xu, D.Y., Read, J.C., Sim, G., McManus, B. & Qualter, P. (2009). Children and 'smart' technologies: can children's experiences be interpreted and coded?. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (BCS-HCI '09). British Computer Society, Swinton, UK, UK, 224-231.

Zaman, B. (2009). *Introduction and validation of a pairwise comparison scale for UX evaluations and benchmarking with preschoolers*. In *Interact*. 2009. Uppsala: Springer.

Zaman, B. and Abeele, V.V. (2010). *Laddering with Young Children in User Experence Evaluations: Theoretical Groundings and a Practical Case*. in *IDC 2010*. Barcelona: ACM.