

Establishing Usability for Interactive Music Applications that Use Embodied Mediation Technology.

Alexander Deweppe,^{*1} Marc Leman,^{*2} Micheline Lesaffre^{*3}

**IPEM (Institute for Psychoacoustics and Electronic Music), Ghent University, Belgium.*

[1Alexander.Deweppe@UGent.be](mailto:Alexander.Deweppe@UGent.be), [2Marc.Leman@UGent.be](mailto:Marc.Leman@UGent.be), [3Micheline.Lesaffre@UGent.be](mailto:Micheline.Lesaffre@UGent.be)

ABSTRACT

This paper proposes a research method to address usability issues in the early stage of development of interactive or collaborative embodied music applications. The central topic is how concepts of usability testing, goal composition, game enjoyment and technology assessment can be actively incorporated in the development and improvement of music applications that use embodied mediation technology. The method presented involves discourse analysis of interviews to develop an approach to constructively use the feedback of test persons to enhance interactive music applications. This analysis is used to determine whether users can properly interact with the system, whether they understand it and how they evaluate the application. The research presented investigates how this feedback can be integrated in a user-oriented development strategy, intended to facilitate the process of developing user-friendlier embodied music mediation technologies. Moreover, the method may be useful in the process of taking interactive music applications out of the research environment and making the transition into existing cultural/musical contexts.

Keywords

Embodied Music Interaction, Music Mediation Technology, User-oriented Approach, Discourse Analytic Method, Music and Personality

I. INTRODUCTION

The process of creating content for an interactive application is often an exercise in obtaining a balance between the inspiration and the technical resources and limitations of the research team working on it. In practice, a music application is often created by a number of researchers who between themselves decide upon the constituent parts of that application, like interfaces that will be used, mappings, feedback, etc. (Wanderley & Orio, 2002).

Using user-related information and feedback is not new to this type of development-process (Stowell, Plumbley & Bryan-Kinns, 2008), but to date it has not been integrated to its full potential into the domain of systematic musicology. To have prospective users evaluate an application and collect their feedback is a procedure often practiced only after the main developmental work is finished. The collected feedback is then only important to make minor alterations to the already existing design.

Given the flexible nature of embodied music mediation technology (Leman, 2008), this presents both tremendous opportunities as well as equally large problems. A broad view of targets have been defined in ‘A roadmap for Sound and Music Computing’ (S2S², 2007) but the importance of involving prospective users in the process of tool building is not stressed. There is a wide range of conceivable applications to choose from (Blaine & Fels, 2003) and a broad spectrum of

cultural environments where embodied music mediation technology could be introduced. Even so, at present, an outspoken goal of how to embed this technology in the existing cultural landscape is lacking.

In this study, the problem at hand is addressed by focussing on the users who under different conditions and on different locations played the social music game ‘*Sync in Team*’, based on the ‘*Musical Synchrotron*’ (Demey, Leman, De Bruyn, Bossuyt & Vanfleteren, 2008), which was also developed at IPEM. More specifically our research aims at encompassing how prospective users want to use this type of technology, what they want it to do and what they like and don’t like about it. This way, we want to 1) evaluate the qualities of the game, 2) gain insight in what shortcomings participants feel the setup has that need to be improved and 3) gather feedback on changes that are made because of this. The method applied to record these user-evaluations, was a structured interview technique that relies on a pre-established set of questions. A random set of people who played the prototype of the music game, were requested to do an interview. These interviews were recorded on video, so that the entire feedback could be processed afterwards.

The objective of this study is to establish how well we can record and apply this feedback. We investigate if and how well discourse analysis (Talja, 1999) of interviews is suited to obtain relevant user-feedback of the participants of this music game. We evaluate this approach and discuss what other user-related methods we should merge it with to obtain a strategy that will allow us to proceed with embedding mediation technology based on embodied music cognition in existing cultural contexts.

II. INTERVIEW ANALYSIS

A. Setup Description

The setup (Figure 1) that was used was a collaborative music game, in which participants were divided into two competing teams of two players. By dancing synchronously with the musical beat and with their team partner, they controlled a visual feedback, projected on the ground. The participants wore a belt onto which a wireless motion sensor was attached. Before the game started, participants were instructed to dance to the beat of the music and simultaneously try to synchronize with the other team member. Every time the game was played, five forty-second music excerpts were played as the musical stimulus the participants had to synchronize with. In the basic setup, the five excerpts were taken from relatively unknown electronic music songs. More music sets were introduced to the game in the following setups. The acceleration of each participant’s dancing was measured with a sensor they wore on the hip. These measurements controlled the visualization of the synchronization level. To determine team players’

synchronicity, the tempo was extracted using an FFT and was compared with the BPM of the music excerpts. The extracted tempos were evaluated and resulted in an increase (when synchronized) or decrease (when not synchronized) of the global score of a team. That score was visualized by the projection on the dance floor (Kaiser, Ekblad & Broiling, 2007). The visual feedback (Figure 2) consisted of two colored dots. Whenever a team performed well, the colored dot of that team expanded. When a team did not synchronize well to the music and/or to one another, the dot that represented their score decreased to its original size and shape. So the precision they performed the set task with, translated into the visualization. The team that performed the task best over a short span of time, got the bigger visualization and the team that performed the task best throughout the entire game, won.

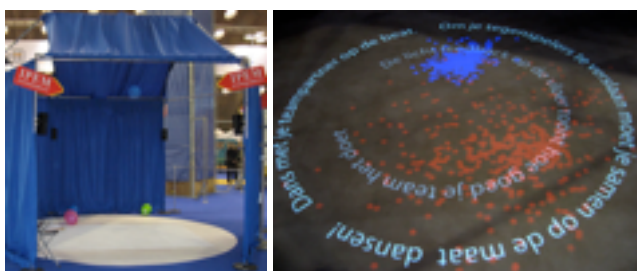


Figure 1. The experimental setting as it was presented at the 2008 science fair in Mechelen.

Figure 2. Close-up example of the visualization presented to the participants.

B. Experiment locations and setup modifications

The experiments the music game was tested in, were executed in four stages and four different settings, which also implies four different interview situations. Each experiment had the same basic components as described earlier, although the particular setting and the state of the art of the development of the game yielded slightly different conditions.

A first series of experiments was done in lab setting at Ghent University, over the course of three days. A basic structure of the game with two conditions was used and evaluated. First, in an individual condition (De Bruyn, Leman, Demey, Desmet & Moelants, 2008), the teams were given the chance to experiment in the music game with their teammate and in the second social condition, the teams played in competition with each other. In this early developmental stage, the visual feedback was projected on the floor and the dots representing the teams remained stationary. Also, only the aforementioned set of five musical excerpts was used in both the individual and the social session. Nearly all participants for this series of experiments were university students.

For the second series of experiments, the game was moved to the Institute of Broadband Technology (IBBT) in Ghent for a public event, an open house day. The introduction round (i.e. individual condition) was omitted from the course of the experiment/game, the visualization was projected on the floor in its definitive, dynamic form and the participants were given the opportunity to choose from different sets of musical

excerpts. The sets the participants could choose from were electronic music, Latin, classical music, hip hop, R'nB and rock music. Their choice could be made either to improve their own score or to negatively influence the performance of the other team. For this second series of experiments, every game contained two sets of five music excerpts in the genres chosen by the two teams and presented to them in a random order. The length of the excerpts was reduced to approximately thirty seconds per excerpt. This series of experiments had the largest demographical variation, as the participants were recruited from the visitors of the open house day.

The third series of experiments took place at a three-day science fair, (i.e. Wetenschapsfeest) biannually organized in Mechelen by the Flemish Government. This fair typically targets children and adolescents, but also adults (parents). For this occasion, a set of children-songs was added to the list of music categories to choose from and the number of sets was doubled, so that the teams could also choose to compete in the same music style. Aside from that, no fundamental changes were made to the setup of the game (Figure 3).

The last series of experiments was done at NEXT08, the two-day national gaming convention in Brussels. In that setup the feedback was projected on a screen, and more music sets were made available for the participants to choose from.



Figure 3. Participants playing the 'synch in team' game at the 2008 science fair in Mechelen.

C. Interview situations

Given the different types of locations (e.g. lab versus semi-public and full-public places), efforts were done to stay as close as possible to the basic predetermined structure of the interviews.

For the first interviews, carried out at the IPeM lab, we had the opportunity to interview all of the participants individually

and with a more semi-structured approach, so that it was possible to elaborate and ask follow-up questions based on some of the answers given by the respondents.

During the second stage, that is the experiments at the IBBT, a random selection of both winning and losing dance-couples was invited to give feedback on their experience with the game and on their performance. Due to time limitations, teams were interviewed instead of individual participants. Therefore, it was necessary to more rigidly adhere to the predetermined interview structure.

The interviews at the third location, the Science Fair, were done in a manner similar to the one utilized at the IBBT, with the only exception that large percentages of the public (and by consequence of the participants) were children and adolescents.

At NEXT08, the noisy environment made it impossible to conduct interviews. Some participants did evaluate changes made to the game and setup, but this feedback could not be recorded.

III. DATA AND RESULTS

D. Interview Data

The central objectives of the questions asked concerned the users' experience of affordance (Gibson, 1979), flow (Csikszentmihalyi, 1990) and presence (Biocca, Jin & Choi, 2001). Specifically for this game, questions were incorporated to investigate the usability (Nielssen, 1994 & Bevan, Kirakowski & Maissel, 1991), elements of goal composition (Nielsen, 1994) and game enjoyment (Ijsselstein et al, 2008). The questionnaire was designed to obtain as much nuanced information as possible: questions were open-ended and a 'yes' or 'no' answer was not an option.

A first section investigated how clear the objective of the game was to the participants and how successful their used strategy was. To establish how well the participants understood the set task, they were asked to describe the task they had to perform to achieve the goal of the game. A second set of questions investigated how hard the participants had to concentrate to perform the task and how they experienced control over the visualization. In the third section of the questionnaire, participants were asked to evaluate the musical and visual features of the game both aesthetically and as a means to play the game. Finally, the participants were asked whether they could think of suggestions, changes or additions that would make the game more appealing to them.

Additional questions dealt with the creativity in the performed task, what type of movements achieved the best results, if there was an evolution in their performance throughout the game, what elements the participants focused on most, what strategies were used to improve the team's performance and eliminate the competition. Finally some questions were asked about the making and correcting of errors, the challenge the participants experienced and the gratification (or frustration) they felt after playing the game.

Some of these questions had to be taken out of the interview structure after the first series of interviews due to time restraints; others were dropped because the response was already implicitly present in the answers provided to some of the other questions.

The video-registration files of the first three locations were processed into accurate transcriptions of the interviews (although at times the answer of the respondent had to be paraphrased). Quantification of the answers that were recorded for the different questions provided the values for the statistical analysis. When respondents gave multiple answers to one question, as was sometimes the case, the most spontaneous answer (usually the first answer) the respondents had provided, was used.

E. Analysis and Results

Forty-five interviews were carried out at the first three locations, of which two could not be used due to technical difficulties with the recording equipment or problems with the execution of the experiment, so the discourse analysis was performed on forty-three interviews. Approximately one third of the interviews were done at each of the locations and thirty-five percent of the interviews were done with individual participants, the others were team interviews. The team interviews were treated in the same way as the individual ones. Of all the respondents half had won the game, half had lost.

In ninety-five percent of the interviews, the goal of the game was said to be (very) clear, and this was confirmed by the results of letting the participants describe the task they had to perform (although during the interviews, sometimes the task was confounded with the used strategy).

In only twenty percent of the cases, respondents stated to have needed very little attention to perform the relatively easy task and for nearly half (46,5%) of the interviewed, performing the task required their constant attention. In approximately seventy percent of the interviews, respondents experienced a good control most of the time over their visualization, whereas twenty-three percent experienced little and seven percent of the participants experienced no control at all. A chi square-test showed a significant relationship between whether the participants had experienced control over the visualization and whether they had won or lost. The explanation for this could be that, although most subjects understood the goal of the game and the set task rather well, they were not able to perform the task properly, hence did not get a good visualization.

All of the interviewed stated to have enjoyed the visualization, so it was not possible to see a rise in the appreciation by changing the visualization from stationary to orbiting. The participants were already enthusiastic about this aspect of the game in the first stage of the development, and in the later stages, they had no grounds for comparison. However, in nearly a third of the cases, respondents did not find the visualization sufficiently clear as a feedback on their performance. In three of the interviews, the respondents suggested to incorporate an actual score into the game.

From the evaluation of the changes made to the game setup based on the participants' feedback, we could see a significant relation between introducing choice of music genres into the game and the appreciation for the music to which subjects had to dance. This relation could not be demonstrated between the three different setups (with the increased choice of music between the second and the third setup). The question whether the music was well suited as a tool to perform the task, did not provide conclusive data either; the reason for this could be

either related to the fact that the participants had to agree upon a certain music style to play the game as a team or the fact the teams also had to compete in the music style the other team had chosen.

IV. EVALUATION AND DISCUSSION

The feedback we recorded allowed us to further improve the setup and to verify whether the changes that were made, had improved the game or not. This process can easily be reiterated in a structured way and with the incorporation of some of the suggested changes and additions to the setup, it can result in a more user-friendly setup and a more enjoyable game.

More importantly, these suggestions give us a clear indication of what participants would like this technology to do. In this particular case, even though the respondents were generally quite enthusiastic about the game, in all but fourteen percent of the interviews participants did suggest changes that would make the game more enjoyable to them. Only a quarter of the respondents could not think of any further additions that could be made to further improve the game. In addition to this, only twelve different answers were recorded to both these questions, which implies that large sections of the interviewed had the same view on how the game should be improved. In a third of the interviews, respondents confided that they would like to play this game with more (teams of) contestants (Feldmeier & Paradiso, 2007). In more than thirty percent of the interviews the suggestion was made to incorporate more (realistic) light-effects in the setup.

These answers do not only give insight into on how participants evaluated the game, but they also indicate that, as a number of participants apparently did not perceive the experience to be realistic enough, what environment the game should be set in.

Although the results of this approach still only apply to the actual setup, they give a good idea of what could make the game more enjoyable and realistic. Overall, using interviews to record feedback about a given application is a quite cost-effective way to pinpoint what things that are lacking from it and early on in the development find out what changes and additions should be taken into account for improvement.

However, for researchers working on the hard- and software development of embodied music mediation technology, the presented approach can hardly be considered a solid foundation. The main shortcoming with using it as a developmental guideline is that it only allows evaluations to be made on existing models and therefore it cannot take prospective participants and possible implementation context into account. And this cannot be addressed solely by going through extensive usability testing.

To work on useful developmental strategies for embodied music mediation technology, a good deal of a priori knowledge about future users (Rentfrow & Gosling, 2003), the context in which they will be used and the purpose they should serve, is vital. Therefore a complementary method is needed to base developmental strategies on to provide this information. For this purpose an online-survey was developed that is still taking place. It uses the open source LimeSurvey-system (<http://www.limesurvey.org/>) to gather precisely that information of a wide range of possible prospective users. Therefore, it is spread with the aid of a number of different

organizations active in the cultural landscape of Flanders, with the goal of screening and recruiting diverse culture-enthusiasts with different musical and cultural preferences for experiments. Although this recruitment approach eliminates the randomness of the participants (as presented in this paper), it provides a far better way to investigate how a specific public likes to interact with a certain type of technology within a given context.

V. CONCLUSION AND OUTLOOK

At present, it is too early to thoroughly evaluate this methodology. However, as the results of the interview analysis prove to be valuable and can be put to good use, we are confident that combining it with profiling prospective users and possible contexts to embed them in, will benefit and accelerate future development of embodied music applications. Development of context-dependent applications cannot dispense with knowledge about prospective users. Combining this user- and context-oriented development with exhaustive usability testing by well documented, specially selected users and incorporating their feedback into the development process is a laborious approach, which is however likely to deliver applicable findings. Therefore, it is highly probable that this strategy will result in a number of plausible and realistic scenarios that will allow the cultural embedding and ensure the ecological viability of embodied music applications.

ACKNOWLEDGMENTS

Thanks to Bärbara Sarmentero, Michiel Demey and Pieter Coussement for the development of and the additional adjustments to the music game 'Synch in Team'. Thanks to Frank Desmet for the help on the statistical analysis. Also thanks to Ivan Schepers for the technical support and the construction of the game stage. Finally, thanks to Nuno Diniz for the technical support on the Limesurvey-system.

REFERENCES

- Bevan, N., Kirakowski, J., & Maissel, J. (1991). What is Usability? *Proceedings of the 4th International Conference on Human Computer Interaction*, Stuttgart, Germany.
- Biocca, F., Jin, K., & Choi, Y. (2001). Visual Touch in Virtual Environments: An Exploratory Study of Presence, Multimodal Interfaces, and Cross-Modal Sensory Illusions. *Presence: Teleoperators and Virtual Environments Vol. 10/3*, pp. 247-265.
- Blaine, T., Fels, S. (2003). Contexts of Collaborative Musical Experiences. *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03)*, Montreal, Canada.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.
- De Bruyn L., Leman M., Demey M., Desmet F., & Moelants D. (2008). Measuring and quantifying the impact of social interaction on listeners' movements to music. *Proceedings of the 5th International Symposium on Computer Music Modeling and Retrieval*, Copenhagen, Denmark.
- Demey, M., Leman, M., De Bruyn, L., Bossuyt, F., & Vanfleteren, J. (2008). The Musical Synchrotron: Using Wireless Motion sensors to study how social interaction affects synchronization with musical tempo. *Proceedings of the 2008 Conference on New Interfaces for Musical Expression (NIME-08)*, Genova, Italy.
- Feldmeier, M. & Paradiso, J. (2007). An Interactive Music Environment for Large Groups with Giveaway Wireless Motion Sensors. *Computer Music Journal, Vol 31/1*, pp. 50-67. Massachusetts: MIT Press.
- Gibson, J. (1979). *The Ecological Approach to Visual Perception*. New Jersey: Lawrence Erlbaum Associates.
- Ijsselstein, W., Van den Hoogen, W., Klimmt, C., De Kort, Y., Lindley C., Mathiak K., Poels, K., Ravaja, N., Turpeinen, M., & Vorderer, P. (2008).

- Measuring the Experience of Digital Game Enjoyment. *Proceedings of the 2008 Measuring Behavior Conference*. Maastricht, The Netherlands.
- Kaiser, G., Ekblad, G., & Broling, L. (2007) Audience Participation in a Dance-Club Context: Design of a System for Collaborative Creation of Visuals. *Design Inquiries*. [online: www.nordes.org].
- Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge, Massachusetts: MIT Press.
- Nielsen, J. (1994) Usability Inspection Methods. *Proceedings of the 1994 CHI Conference*. USA, Boston
- Nielsen, J. (1994). Goal Composition: Extending Task Analysis to Predict Things People May Want to Do. [online: <http://www.useit.com/papers/goalcomposition.html>].
- Rentfrow, P., & Gosling, S. (2003). The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences. *Journal of Personality and Social Psychology*. Vol. 84/6. pp. 1236-1256.
- S2S² Consortium. (2007). *A Roadmap for Sound and Music Computing*.
- Stowell, D., Plumbley, M., & Bryan-Kinns, N. (2008). Discourse Analysis evaluation method for expressive musical interfaces. *Proceedings of the 2008 Conference on New Interfaces for Musical Expression (NIME-08)*, Genova, Italy.
- Talja, S. (1999). Analyzing Qualitative Interview Data: The Discourse Analytic Method.
- Wanderley, M., & Orio N. (2002) Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal*. Vol 26/3 pp. 62-76. Massachusetts: MIT Press.