



Comparing Computer Versus Human Data Collection Methods for Public Usability Evaluations of a Tactile-Audio Display

Maria Karam

Centre for Learning
Technologies,
Ryerson University
Toronto, Ontario
Canada
maria.karam@ryerson.ca

Carmen Branje

Centre for Learning
Technologies Toronto, Ontario
Canada
cbranje@ryerson.ca

John-Patrick Udo

Centre for Learning
Technologies,
Ryerson University
Toronto, Ontario
Canada
judo@ryerson.ca

Frank Russo

Professor,
Department of Psychology
Ryerson University
Victoria St.
Toronto, Ontario
Canada
russo@ryerson.ca

Deborah I. Fels

Professor and Director,
Centre for Learning
Technologies
Toronto, Ontario
Canada
dfels@ryerson.ca

Abstract

We present a public usability study that provides preliminary results on the effectiveness of a universally designed system that conveys music and other sounds into tactile sensations. The system was displayed at a public science museum as part of a larger multimedia exhibit aimed at presenting a youths' perspective on global warming and the environment. We compare two approaches to gathering user feedback about the system in a study that we conducted to assess user responses to the inclusion of a tactile display within the larger audio-visual exhibit; in one version, a human researcher administered the study and in the other version a touch screen computer was used to obtain responses. Both approaches were used to explore the public's basic understanding of the tactile display within the context of the larger exhibit.

The two methods yielded very similar responses from participants; however, our comparison of the two techniques revealed that there were subtle differences overall. In this paper, we compare the two study techniques for their value in providing access to public usability data for assessing universally designed interactive systems. We present both sets of results, with a cost benefit analysis of using each in the context of public usability tests for universal design.

Keywords

Universal design, usability studies, public usability, crossmodal displays, sensory substitution, audio-tactile interfaces, evaluation techniques

Introduction

Designing and implementing an interactive system that offers universal access to music, sound, and entertainment to a diverse range of users poses many challenges. Once such a system is designed, however, it must be evaluated by as wide a range of user groups as possible in order to determine whether it actually achieves the goal of broad access. While laboratory studies have served as one of the most effective methods for assessing interactive systems with users, it can be difficult to access diverse populations of potential users. In addition, laboratory studies are often contrived and controlled, providing limited ecological or in situ validity, which may be highly desirable for a system that provides access to entertainment for the public.

The Emoti-Chair is a system that aims to provide universal access to the emotional expressiveness that music and sound can add to the entertainment value of a visual presentation for users (Branje, Karam, Russo, & Fels, 2009; Karam, Russo, & Fels, 2009). It uses a variety of tactile displays to create crossmodal representations of the emotional content expressed through the audio stimuli that accompanies visual media such as film or television. One important aspect of developing the Emoti-Chair is evaluating its usability and entertainment value in an appropriate venue where a variety of users can experience it.

Relying solely on laboratory experiments to evaluate the entertainment value or comfort of a system like the Emoti-Chair is problematic in that laboratory environments are not typically conducive to supporting relaxed, fun experiences for users. It is also difficult to attract a large and diverse population of potential users, including those with different hearing abilities, to a laboratory environment. Thus, real-world feedback from many potential users of the system may never be obtained during the development lifecycle. Also, it is often the case that systems that fare well in laboratory studies end up as commercial products without ever having been effectively evaluated in the actual interaction context they were designed to support. This is even more problematic for systems based on universal design principles (see Burgstahler, 2008 for a review), because there are so many potential user groups that must be accounted for in the design.

Testing interactive systems within a public space such as a public museum offers an alternative method for retrieving information from a wide range of users and their interpretation of the functionality and usability of a system within the context of the intended interactions, which can lead to improved usability studies in general (Christou, Lai-Chong Law, Green, & Hornbaek, 2009). While testing in a real-world setting can yield a richer set of user feedback, there are also many obstacles that must be addressed before new interactive systems can effectively be presented in a public space even for evaluation. These include ensuring that the system is robust, safe, reliable, and effective for use in an uncontrolled public environment. In addition, automatic data collection methods can be deployed to leverage the large numbers of potential users who visit public spaces to supplement the data that can be collected by human researchers who may not be available at all times.

In this paper, we present a report on a public usability study that was conducted at a public science museum with the Emoti-Chair. The version of the Emoti-Chair used in this study was extensively reworked to prepare it for display as the tactile component of an audio-visual (AV) presentation targeting global warming and the environment for high school children from the northern territories of Canada, and from a more urban setting within Toronto. We wanted to determine if the role of the chair in the exhibit was apparent to users, while evaluating the overall comfort, clarity, and context of the crossmodal displays for those users. We report on the efforts taken to prepare the chair for public display and on the two approaches (automatic and human facilitated methods of data collection) used to evaluate the effectiveness of the chair.

Usability for Universal Design

Universal design differs from conventional design practices in that designers are asked to anticipate the varieties of users and uses for their product that they may not have previously realized and/or discounted. Whereas conventional design practices encourage designers to address the needs of the anomalous average user, progressive designers who adhere to the principles of universal design realize that this mainstream user simply does not exist (Udo & Fels, 2009). As such, these designers understand that users have needs and preferences that

differ, not only between subjects, but also over time (Rose & Meyer, 2002). Universal design theory proposes that a “user population” and “potential uses” of designs must be very broadly defined to include a variety of users and user abilities that would not be considered within the definition of average. Shneiderman (2000) argues that it is the responsibility of technology enthusiasts to “broaden participation and reach these forgotten users by providing useful and useable services” (p.87). In designing technology that is proactively built to ensure use by individuals with a wide variety of needs, designers give all users the ability to customize and individualize the way in which they use technology.

The Emoti-Chair was developed as an alternative avenue for translating audio stimuli, serving to simultaneously substitute or complement sound stimuli with movement and other tactile sensations. Musicians and artists worked with students to create an audiovisual entertainment experience for the Emoti-Chair that was indicative of the children’s understanding of rural versus urban spaces and the role of technology in society. Students who designed the original audiovisual presentations were given the opportunity to compose a tactile experience for the Emoti-Chair. The design process used to assemble the exhibit held true to many of the tenets of universal design, mainly in the inclusion of the original designers (the children) throughout all stages of the creation of accessible media. In addition, no special enhancements or interface alternatives (assistive technology) were required to enable people who were deaf or hard of hearing to use the system. The same access was provided to all users, regardless of their ability.

The chair was constructed to take into account as many of the principals of universal design as possible, although there were some shortcomings, which will be addressed in further research. For example, the design did not address the needs of wheelchair users who may not wish transfer from their own chair to the Emoti-Chair.

The Emoti-Chair

The Emoti-Chair was designed to present sound stimuli as tactile sensations to the body. Originally developed to assist deaf and hard of hearing (HoH) people in accessing sound information from films, the Emoti-Chair system is based on universal design principles, which aim to increase access to all forms of media for users of all abilities. The system is installed in a large comfortable gaming chair, which has been augmented with a variety of tactile displays aimed at communicating different elements of sound, including speech prosody, music, and background or environmental sounds as physical stimuli. These displays include motion actuators, air jets, and vibrotactile devices that are distributed throughout the chair. Each of these devices can be independently controlled by a computer system, offering the potential to develop a lexicon of tactile sensations that can represent audio events occurring in music, film audio, or live performances.

One of the theoretical contributions of the Emoti-Chair is the model human cochlea (MHC) (Karam, Russo, Branje, Price, & Fels, 2008). The MHC is a sensory substitution technique we use to present the musical component of sound as vibrations. The MHC distributes sound to arrays of vibrotactile channels that are embedded into the back, seat, and arms of the Emoti-Chair (see Figure 1). Our approach attempts to leverage the skin as an input channel for receiving musical signals and other sounds as physical vibrations. We placed the signals relating to music along the back and seat while signals representing human speech are presented along the arms of the chair. Background noises that are not related to music or speech are communicated using motion actuators and air gusts, which enable us to explore the sensory substitution of different types of sounds as unique tactile sensations. The Emoti-Chair uses the air jets and the physical actuators to represent other elements of sound such as wind blowing, earth shaking, or subways moving, although the mapping of sensations to sound is a topic for further investigation.

While this chair is primarily intended for use as an entertainment device, an additional function is to provide a platform for supporting research into the use of different tactile sensations as substitutions for sound. Through the different tactile elements included in the Emoti-Chair, it is possible to conduct investigations that can contribute to our understanding of the relationships between sound and touch. Each of the tactile devices is fully configurable and can be customized as required for the given application.

Furthermore, we are investigating the chair as an artistic contribution that can add a third modality, the sense of touch, to expressive art interactions that transform audio-visual (AV) displays into tactile-audio-visual (TAV) displays.

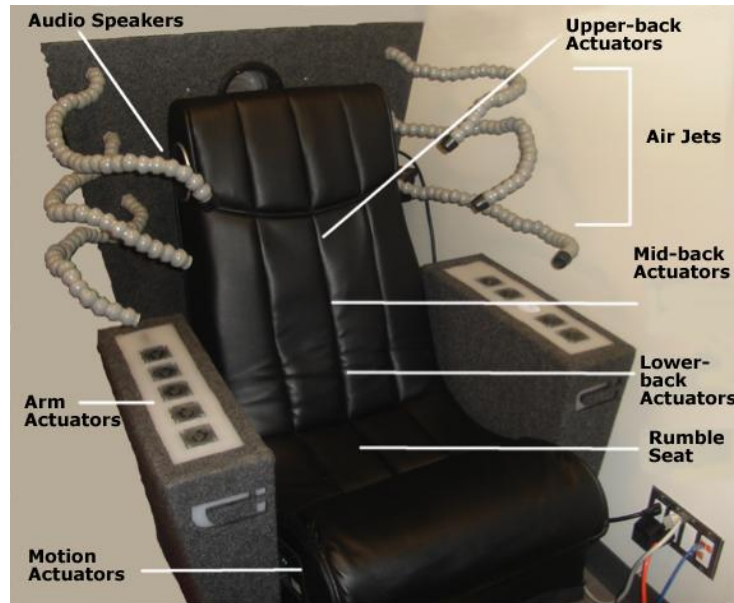


Figure 1. Emoti-Chair that was used in the museum exhibit.

Physical Preparation

To begin this study, several requirements had to be met to ensure that this novel interface would be appropriately integrated into the exhibit. First, because the chair would be exhibited in a public domain, it was necessary to reinforce the prototype to handle the volume of users who would be seated in the chair over the four months of the exhibit, which reached over 6,000 by the second month. This involved reinforcing and refining the prototype to ensure that it would be functionally and physically robust for the duration of the exhibit. The body of the chair was reinforced using a wooden structure. In addition, the hardware was upgraded and modified to account for the continuous use it would have to endure. The peripheral devices, including a large screen display, air supply, and Internet connection, were integrated into the infrastructure of the museum. All potential pinch points were eliminated and electronics access points were locked to minimize the safety risks.

Technical Preparation

One of the issues addressed for this public display was the need for remote administration and maintenance of the system. Because the Emoti-Chair was a research prototype, we had to account for potential support issues that could arise from extensive unsupervised use. Technical adjustments that were outside of the duties of the museum staff had to be handled remotely by our research staff. As such, our researchers ensured that they were able to remotely access the system and, in cases of emergency, log in and reconfigure the system in the event of a failure or crash. The entire system also underwent extensive testing and modification to ensure robustness, safety, and ease of administration in the event of any technical problems that occurred when our research staff were not on site. These problems included unexpected software and hardware crashes or issues that could potentially arise from such a high volume of users.

We note that an incidental benefit of conducting research in a public domain was the level of stress testing that the system received. To support the museum staff in administering the Emoti-Chair, we modified our software control interface to provide a large on-off switch that museum administrators could access to start or stop the display. Otherwise, the system was fully automated to run continuously and in synchronization with the AV component of the

display. An infrared switch was placed on the chair to enable the system to be activated only when a visitor was seated. Because we were working with an early prototype, we anticipated that many problems would be revealed over the course of the exhibit. Our research staff could be contacted by the museum staff, who would call in the event of a system failure for remote administration of the system. Although the system was easy to use, the user interface and overall components were not fully developed for public use, and our research team spent many hours monitoring the system to ensure that the display was running as designed.

Content Production

The content creators were given an introduction to the functionality of the chair and its tactile sensations so that they would be able to create their part of the exhibit and incorporate the tactile sensations of the chair. This was an interesting process because most of the people on the content creation project did not have any prior experience with tactile displays or with the chair itself. We held a series of introduction and familiarization sessions with the contributors, which included high school students, musicians, teachers, and researchers. Everyone on the project had the chance to operate and experience all of the tactile elements of the chair, and each contributor was given a brief set of guidelines to use when integrating sound and video with the tactile display. This represented a creative component of the Emoti-Chair, which was a new experience for the AV producers collaborating on the project.

One of the challenges of this study was to determine how AV content providers would approach the inclusion of the tactile components in their presentations. Because we had a short period of time to introduce the content producers to the system, we used an ethnographical approach that enabled us to observe how the content producers would use the tactile display components and incorporate them into their work. The researchers on the project also benefited from this process as we had the opportunity to obtain firsthand experiences of how artists would approach the design of a TAV display. We considered several different approaches for designing the tactile element of the display, and we discuss some of these next.

Designing audio-visual-tactile content

Most of the people working on the project were very excited about the possibility of appealing to the tactile senses in addition to audio and visual; however, it was not clear how they would approach this as a design problem. To bootstrap the process, we suggested that they consider the energy usage as a metaphor for assigning tactile sensations to their AV content. Some of the students determined that the more energy use that was taking place in the video would be reflected by a stronger set of tactile sensations, while others were not sure how to map any of the tactile sensations onto the AV content of their exhibit. For the music displays provided in the arms and back of the chair, there was little confusion. Students simply directed the music component of their presentation to the tactile display. To assist students in creating the TAV presentation, we developed a preliminary set of guidelines to facilitate the design of tactile sensations by mapping them onto the visual events in the film, as described below:

- Ensure each repeating event in the film corresponds to the same tactile sensations.
- Ensure that the different levels of tactile sensations correspond to the intensity of the associated event in the AV presentation.
- Reuse of tactile events should not conflict with existing mappings of tactile sensations onto AV events.
- Similar levels of sensations for each of the tactile displays should be applied to events with similar contexts.
- Musical content should be displayed through the MHC embedded into the back of the chair.
- Speech audio should be displayed in the MHC embedded into the arms of the chair.
- Intensity levels of the different sensations should correspond to the intensity of the events they are used to express.

Once the chair was prepared and the tactile content was choreographed, we designed the public usability study that would take place during the exhibit to gather initial responses from the public about this type of multimodal display. A picture of the chair as installed at the museum is

shown in Figure 2. The large screen display was mounted from the ceiling and placed in front of the chair where it would be visible.



Figure 2. Children trying the Emoti-Chair at the Ontario Science Centre.

The Study

The study we conducted focused on several aspects of the chair and its associated tactile sensations. These included learning whether visitors to the exhibit noticed and comprehended the relationship between the exhibit and the tactile sensations of the chair and exploring which of the tactile devices were most appealing to different users. We also wanted to explore the automation of data collection as an alternative to having only human researchers administering the questionnaires. We were interested in determining whether there were any significant differences in the two data collection approaches to obtaining user feedback on the system, towards further developing a methodology for conducting public usability evaluations.

Additional criteria for designing the questionnaires included limiting the amount of time and effort that was required from the museum visitors, so we only focused on the most critical aspects of the exhibit and the TAV for this study, and rewording the questions to be accessible to a young child's reading ability.

Questionnaire Design

The key features of the display we hoped to examine through the questionnaire included understanding whether users could perceive the relationships that linked the tactile, visual, and audio elements of the display with the exhibit, assessing the level of enjoyment users derived from the experience, and the level of comfort the chair would provide. We modified an existing questionnaire used for in-laboratory evaluations (Karam et al., 2008) to fit in with the exhibit and the issues being explored in the public usability environment.

As this was our first attempt at conducting public usability studies, we conducted several iterations of pilot studies that led to the final version. This study was primarily aimed at hearing individuals who could provide us with a baseline understanding of the connections that could be drawn between the sound and the tactile sensations. There were a total of 25 questions in each questionnaire.

The questionnaires asked users to indicate their sex, age group (under 10; youth 11-19; adult 20-64; senior, over 65), which presentation they watched, to rate their comfort and enjoyment of the chair, the relationship between the chair and the exhibit, and the chair's fit with the

exhibit. We also asked about the amount of background information participants read before entering the exhibit, as well as preference and enjoyment ratings for each of the tactile devices.

Two methods were used to collect questionnaire data from museum patrons: researcher administered and computer administered.

Researcher-Administered Questionnaire

The researcher-administered questionnaire presented the same questions as the electronic version. Researchers wore badges to indicate their name and university affiliation. After a participant experienced the Emoti-Chair, a researcher would approach them to see if they would be willing to answer a few questions to provide feedback on the exhibit and help us improve the design of the system.

We interviewed participants of all ages, but asked permission before approaching children. Most of the participants were eager to take part in the study; this may have been due to the type of museum (science oriented), where people were pleased to contribute their feedback towards improving the technology. Researchers were present at the museum three days a week for a period of two months and obtained a large number of responses; however, many of the responses were gained during the pilot study and were not included in our analysis.

Self-Administered Questionnaire

The self-administered questionnaire was presented using a PC computer and 15-inch touch screen monitor. The screen was mounted on the wall located behind the Emoti-Chair installation. The self-administered questionnaire was identical to the research-administered questionnaire, but the format was modified for the touch screen. The questionnaire application was developed in C# using Microsoft Visual Studio 2008 and a MySQL relational database. Each question was presented on a single screen, with the question displayed in large text at the top of the page and a large button below to represent possible responses, as shown in Figure 3. Additions were also made to the digital version of the questionnaire to make it more suitable to the touch screen format and to account for potential problems that could arise in this public domain. For example, the visual layout was designed to ensure that participants could easily distinguish the question text from the answer buttons.

Also, we implemented a monitoring system that could help us determine when there were problems with a particular session. For example, to determine if a user abandoned their questionnaire session, we used a timer to monitor the system at the start of each question. If the timer reached 30 seconds without a response, the questionnaire was reset. The session was terminated and marked as incomplete. If the participant responded within the 30-second timeframe by selecting one of the button responses, the timer was reset to 0 and a new question was displayed on the screen. After each response button was pressed, its value was saved in an SQL database on our remote server, and the screen was updated with the next question. The survey ended when all questions were answered. The computer running the survey was connected to the Internet so that data could be uploaded to our remote server on a regular basis. Maintenance could also be performed remotely when the survey application or computer failed. We also added an introduction screen to explain what the questionnaire was about and to check if participants had already tried the chair. If they did not, they were asked to return and complete the questionnaire only after trying the chair. This was not a problem in the human-administered version.

Participants

The study was focused on evaluating the system with hearing individuals. The participants who took part in the study included children under the age of 10, seniors over age 65, as well as people between these ages. All were visitors to the science museum, and all took part in the study based on their interest in the display.

3. Which audio/video presentation did you watch?

Time remaining to answer:
29

You have answered:
3 of 25 questions

Morning Footprint

Afternoon Footprint

Evening Footprint

Why do Toronto Students Use So Much Energy

I do not remember

Figure 3. Screenshot of one question from the self-administered questionnaire

Results

We report on the results from the researcher-administered and the self-administered questionnaires, in addition to the advantages and disadvantages experienced with both methods. Results from the human-administered questionnaire were described in a previous paper (Branje et al., 2009). In this section, we review results from the self-administered version and then compare these results to those of the human-administered version. A Kolmogorov-Smirnov test for normality showed that the data departed significantly from normality, due to the difference in the number of participants and question responses between the two conditions. As a result, a non-parametric Mann-Whitney test was performed to examine whether there were significant differences between the two survey methods for the respective survey responses. Seven of the eighteen response questions, including four of the five questions concerning the detectability of the various aspects of the Emoti-Chair, were found to have a statistically significant difference ($p < 0.05$) between survey methods (see Table 2).

Researcher-Administered Questionnaire Results

Having a researcher present to administer questions to a participant is an effective way of reducing problems associated with the integrity of the results and for obtaining additional information through open-ended questions and interviews about the user experience with the system. Humans have an advantage over computers to make observations, alter their approach, or answer questions while the participant is engaged. However, it is also more challenging to approach people in a public setting without interfering with the natural flow of the user experience.

Several existing problems common to human-administered questionnaires were also present in this study, including the potential for participants to feel the need to respond positively because they were speaking with a person and may not have wanted to provide negative feedback. While human researchers can also obtain additional information through interviews and open-ended questions, we were more interested in leveraging the large turnover of visitors who could provide us with general feedback on the system. This approach provided us with a broad overview of the users' perspective on the system within the public entertainment domain, towards further identifying problems and other issues with the universality of the design.

Self-Administered Questionnaire Results

Many aspects of the self-administered questionnaire made it preferable over the researcher-administered one. First, although initial programming work was required to develop the automated questionnaire, very little continued labor was required from the research team to administer the questionnaire. The questionnaire was remotely accessible and so data could be collected on a regular basis and any problems with the questionnaire could be identified and fixed early on in the pilot stage of the study. The self-administered questionnaire yielded a high number of participants, and even after anomalous entries were deleted, over 550 usable questionnaire entries were collected over a two-month period.

As we expected, a major problem with the self-administered questionnaire was associated with the public nature of the display. Researchers observed participants pressing question keys in rapid succession and observed others walking away from incomplete questionnaires.

As this was a children's science museum, there were many attention grabbing aspects of the exhibits competing for visitors' attention, and it appeared that in some cases the questionnaire was not able to retain the participants' focus long enough to complete the 25 questions.

Comparing Results Between Survey Methods

In this section, we discuss the differences in survey results between the two questionnaire methods. These are presented and organized based on sections of the questionnaires. Although most of the results showed no statistically significant difference between the two methods, this is an expected result, suggesting that both methods yield similar results. The few that did show difference are discussed below.

Table 1. Age groups of the participants who took part in each survey method.

| Age category | Human survey number | Computer survey number |
|----------------------|----------------------------|-------------------------------|
| Child (10 and under) | 20 (27%) | 110 (20%) |
| Youth (11-19) | 34 (46%) | 242 (44%) |
| Adult (20-64) | 20 (27%) | 177 (32%) |
| Senior (65 and over) | 0 (0%) | 21 (4%) |
| Total N | 74 | 550 |

Demographics

There were no participants in the oldest age category (65+) surveyed using the human-administered questionnaire and only 21 were surveyed using the computer-administered survey. The largest number of participants for both the computer and human administered surveys were youths, followed by adults and children. This seems to reflect the demographic of visitors to the museum, which is geared towards schools where a few adults supervise many children.

The distribution of participants for both methodologies was statistically significant, such that the groups were not equally divided by age. We suspect that the variation in age group taking part in the studies may be due to several factors. First, children under 10 are likely to have lower comprehension and reading levels than youth or adults, which may have discouraged many from taking the self-administered survey. Youths may have been more confident in taking the survey, because they are more likely to be comfortable with computers systems and touch screens. Additionally, the survey was mounted on the wall approximately 36 cm off the floor, which may have been out of reach for some of the younger visitors.

Detection of Tactile Sensations

One of the most important aspects of the Emoti-Chair explored in this study was the detectability of each of its tactile components. We found that responses to questions about the detectability of sensations were significantly different for the two survey methods. Table 2 shows the mean and standard deviation (sd) for questions with significant results. In general, participants interviewed by the researchers indicated that they could detect the whole body vibrations, the lower back vibrations, and arm vibrations more than participants who took the self-administered survey. However, upper back vibrations were found to be significantly less obvious to respondents of the researcher-administered survey than in the automated version. There were several possible reasons for this result. First, the signals emitted through the tactile devices located on the upper back were relatively weaker than those on the lower parts of the back. The upper back stimulators were used to convey high frequency signals, while the seat presented vibrations in the mid to low frequency ranges. At high frequencies the voice coils do not vibrate as strongly as they do for lower frequencies, making them more difficult to sense.

Table 2. Significant differences between computer-based and researcher collected survey data using the Mann-Whitney U test. Ratings were provided on a 5-point Likert scale where 1 is the least detectable and 5 is the most detectable.

| Tactile sensation variable | U | Z | df | Computer-based | | Researcher collected | |
|------------------------------|-------|-------|-----|----------------|------|----------------------|------|
| | | | | M | SD | M | SD |
| Detect upper back sensations | 16335 | -2.99 | 623 | 2.96 | 1.38 | 2.44 | 1.34 |
| Detect lower back sensations | 16757 | -2.69 | 623 | 2.94 | 1.44 | 3.41 | 1.39 |
| Detect arm sensations | 17061 | -2.48 | 623 | 3.10 | 1.41 | 3.52 | 1.31 |
| Detect whole body sensations | 14546 | -4.09 | 622 | 3.17 | 1.34 | 3.85 | 1.08 |

Though we could adjust the power delivered to each of the vibrators individually, this could not be done on an individual basis in this environment. Also, there were fewer participants who took part in the human-administered evaluation, which may have yielded a larger percentage of respondents who did not notice the higher frequencies along the upper back. Finally, people of different stature may have experienced the upper back vibrations along different parts of their back (e.g., taller people may have felt the upper back vibrations in their mid-back section instead), but our data did not address individual body types for this study.

The difference in response to the detectability of the sensations in the middle back was not significant between computer and human versions of the survey. The anonymity offered by the computer-based survey may have elicited these more candid/negative responses, leading to more honest responses from those who did not detect certain sensations. Anonymity or perception of privacy is reported to be one of leading explanations for the increase in honest responses that are associated with computer-based surveys (see Richman, Kiesler, Weisband, & Drasgow, 1999 for a meta-comparison between computer-based and human delivered surveys for non-cognitive type data).

Enjoyment Ratings

Significant differences in responses were also found in the enjoyment ratings for the back vibrations and linear actuators for the two survey methods. Table 3 shows a comparison of means and standard deviations for each survey method for this category of question. Generally, participants responding to the self-administered questionnaire rated these features of the Emoti-Chair higher on enjoyment than did the participants who took part in the researcher-administered questionnaire.

Table 3. Significant differences between computer-based and researcher collected survey data. Ratings were provided on a 5-point Likert scale where 1 is least enjoyable and 5 is most enjoyable.

| Enjoyment variable | U | Z | df | Computer-based | | Researcher collected | |
|--------------------|-------|-------|-----|----------------|------|----------------------|------|
| | | | | M | SD | M | SD |
| Back vibrations | 15284 | -3.73 | 623 | 3.25 | 1.3 | 2.60 | 1.43 |
| Chair rocking | 16889 | -2.61 | 623 | 3.29 | 1.32 | 2.83 | 1.41 |

Other enjoyment variables that were not statistically significant were overall enjoyment, overall comfort, and enjoyment of the air jets, seat vibrations, and arm vibrations.

These results would seem counter intuitive to literature (e.g., Richman et al., 1999) regarding the candidness of ratings in computer-based surveys versus researcher-based surveys as the computer surveys showed a higher enjoyment rating. There could be two possible explanations for this result. The first relates to the possibility that this result is a Type I (a false positive) error in that we compared 18 questions and we would expect that, by chance, at least one question (5%) would be significant. The second explanation is that enjoyment is a more abstract concept compared to whether or not the sensation is detectable. The researcher-based survey allowed elaboration on the concept of enjoyment and that could have led to the more negative ratings. If enjoyment were further explained or better worded in the computer-based survey, a more negative answer may have resulted.

Information and exhibit responses

A significant difference was also shown for participant responses to the two survey methods for the perceived alignment of the Emoti-Chair with the concept of an eco-footprint. Fit was rated on a 5-point Likert scale with 1 being *no alignment* and 5 being *completely aligned*, $U(623) = 17210$, $Z = -2.38$, $p < 0.05$; $M(\text{self administered}) = 2.92$, $SD = 1.39$; $M(\text{researcher administered}) = 2.51$, $SD = 1.18$. Participants responding to the self-administered questionnaire found that the Emoti-Chair fit better with the eco-footprint concept than did the participants responding to the researcher-administered questionnaire. Although, the exact reason for this difference is difficult to ascertain, we speculate that people who took the initiative to complete the self-administered survey may have also spent more time reading through the exhibit material that outlined the purpose of the display.

There were five variables in this category with non-significant differences between survey types: the relationship between the video material and the chair rocking, air jets, background information provided, vibrations, and the fit of the Emoti-Chair with the polar year theme of the exhibit.

Analysis

The results of this study are promising, suggesting that there is value in using an automated response system to obtain user feedback in a public-usability domain. While it was shown that the responses to both surveys yielded similar trends for most of the questions, several interesting findings were revealed, suggesting where improvements to our approach can be made.

First, a more detailed description of the system and the different tactile sensations they provide would be beneficial for those who will be experiencing the system. For example, some people may not have experienced all of the devices in the chair, especially in the case where they were not fully seated and could not feel the back vibrations. This was a problem for both versions of the study, because participants may not have wanted to ask the researcher or re-read the display information to clarify what each of the different sensations were.

Second, the availability of the self-administered questionnaire over the human administered version was advantageous for both researchers and participants. Human researchers cannot always be available to interview participants, while the automated version is always in operation during museum hours. Human researchers may also be selective in who they approach to interview, while participants may also be more hesitant to approach researchers than they would be walking up to the computer.

Third, the inclusion of a touch screen to obtain user feedback adds another element of interest for visitors to the museum, who often seek out new experiences. The computer adds an interesting element for users, who may be keener to explore interactions with a computer screen associated with the exhibit than to break the flow of that experience to speak with a human researcher. However, human interviewers offer opportunities to explain unclear terminology or concepts that are presented to the interviewee. The self-administered version does not offer the same level of communication; however, we believe that a carefully designed help system could facilitate access to information about the system. Humans may also be able to probe interviewees to clarify their answers or to gather further data such as why the

interviewee had a particular response. In addition, interviewers who are frequently at the museum (multiple days and weeks) can specifically target interviewees in missing groups. Although this did not happen in our case (we are missing responses from people in the 65+ group from the human interviewer responses), it is possible to adjust. The self-administered version cannot selectively target respondents from demographic categories with low numbers of participants.

Another limitation of our approach was the lack of data that showed people's levels of understanding of the exhibit content including the Emoti-Chair. This is some of the most difficult information to collect, particularly in a public setting. Using typical multiple-choice, fill-in the blank, matching or essay type questions is problematic as new factors such as difficulties with literacy, test-taking anxieties, compliance with completing the survey, and the relationship between scores on "tests" and understanding are introduced. In addition, people may feel embarrassed about wrong answers or appearing to not have paid attention, which may lead to hesitation when asking for clarification or explanations. In addition, while the goal of our reported study was not to evaluate the participant's level of comprehension of the display, this is a factor we would like to explore in future studies. One approach to this would be to implement a "test your knowledge" element as a separate entity that is integrated with the exhibit. In this particular museum a number of "test your knowledge" stations already exist, which are designed to be fun and engaging. Using the museum's existing approach combined with a tracking mechanism that collects people's responses could provide this type of data.

Although there is considerable research showing that the presence of human interviewers can introduce more response bias (Richman et al., 1999), there are benefits to using both human and computer-generated data collection techniques. The computer version is valuable for acquiring large number of participants, while the human administered versions are useful in testing the questions and in discovering any potential problems with the exhibit, or the questions, before the automated version is finalized. Human collected data are also important in gathering more detailed probing data that explains the user's perception of the exhibit, the system, and the questions.

Recommendations

There are many informative results obtained through this study. We discuss these in order of their appearance in the public usability domain.

System Reinforcement

There are several issues to consider before taking a system out into a public domain. First, it must be at a stage of development where it can be reinforced and safely used by the general public. Some technology may be too fragile or unsafe for this type of environment, and given the extent of the users that can potentially access this technology, safety is a major concern. While there are many systems that have been developed in the research lab, it is essential to assess the feasibility of upgrading and re-enforcing these devices before spending the time and resources to prepare them for a public domain. Before taking a system into the public usability domain, we recommend that feasibility studies or other laboratory work be conducted to ensure the potential of the system for future developments.

Universality Means for All Users

While it is possible to exclude many people from a system that is designed for universal access, conducting a public study may reveal problems that specific user groups may have with the system. In our case, we found that older participants did not enjoy the actuators or the air jets, and while we thought that these were important elements in the translation of sound into tactile sensations, results from our study showed otherwise. One approach to modifying the system without having to completely redesign these features is to offer the option to turn off one or more of the sensations as the user requires.

Ethics Approval

Another interesting aspect of this study involved the ethics application we submitted before conducting this research. Typically, studies involving human participants require approval from the ethics committee of a university. However, in our case, we were only interested in people's feedback to the system, rather than in specifically assessing a user's ability to use software. As

a result, the ethics review board did not require a formal ethics application. However, they asked us to follow good ethical practices; a brief description of the research was provided to all participants including a discussion of voluntary participation and the right to stop the process at any time. In the computer-based survey, a short description was provided at the front end of the survey. Participants could abandon the computer survey at any time and it would simply time out and return to the beginning page.

Additional constraints from our ethics board stated that we could not record the voice or image of any user, and that we must ask permission from parents or guardians before approaching any children under the age of 10.

Given the nature of our study, the ability to obtain user feedback from the public domain greatly increased the opportunity to present our work to a large number of users and to gather valuable feedback for longer periods of time, without having to first obtain written permission. This will be included as part of the methodology we are developing for conducting future usability studies in the public domain.

Conclusions and Future Work

We presented a study that compared two different approaches to gathering user feedback on a system designed to provide universal access to sound information using a tactile display called the Emoti-Chair. The study focused on determining the enjoyment, comfort, and comprehension levels of the tactile displays in the context of a public usability study situated at a science museum. Results from both versions of the study showed no significant differences for most of the questions, which suggest that a computer survey can be effectively automated to acquire basic feedback about the user's experience with the system in the public domain.

The next phase of our public usability study will focus on deaf and hard of hearing individuals who will have the opportunity to use the Emoti-Chair in a different public space, located in a deaf community and cultural centre in Toronto, Ontario. As such, we will further develop our public usability methodology with a more specialized user group that can provide us with a different kind of feedback on the sensory substitution of sound through tactile vibrations for entertainment applications. Results from the study we have reported have led to the development of a simplified version of the Emoti-Chair, which removes the air jets and motion actuators from the chair. This new version is now in place at the Bob Rumball Centre for the Deaf in Toronto, where public usability studies focusing on the deaf and hard of hearing communities are underway.

Practitioner's Take Away

We have found that it is important to consider the following concepts when creating systems using universal design principles:

- Use automated questionnaires to increase participant numbers in evaluations.
- Use public usability studies to supplement lab experiments with real-world data.
- Modify existing evaluation tools to extend into the public domain.
- Include as wide an audience in evaluations to ensure universality of the design.
- Expect to alter major features of the design to ensure target users are addressed.
- Select only technology that is robust enough to withstand constant public use.
- Reinforce and secure systems to ensure safety of users and the equipment.
- Restrict public studies to short, succinct questions and questionnaires to maintain ease of ethics approval, and focus the study on the broader aspects of system interactions.
- Ensure proper assistance is in place to accommodate users with special needs during the study. Sign language or other interpreters, or special access needs are essential to address when conducting studies on universal designs.

Acknowledgements

Funding for this research was provided by an NSERC/Canada Council strategic grant. We gratefully acknowledge the Ontario Science Centre for allowing us to carry out research at their facility. We also thank Kathleen Suarez-Mason, Krystian Wilk, and Pedro Rigo for running the human-administered portion of this study, all of the participants who completed our surveys, and Ana Klasnja, the curator of the Ideas Gallery at the Science Centre where the exhibit was held.

References

- Branje, C., Karam, M., Russo, F., & Fels, D. (2009). Enhancing entertainment through a multimodal chair interface. IEEE Toronto International Conference–Science and Technology for Humanity, Toronto, In press.
- Burgstahler, S. (2008). *Universal Design Process, Principles and Applications*. Retrieved July 15, 2008, from University of Washington:
<http://www.washington.edu/doit/Brochures/Programs/ud.html>
- Christou, G., Lai-Chong Law, E., Green, W., & Hornbaek, K. (2009, April 4-9). Challenges in evaluating usability and user experience of reality-based interaction. In Proceedings of the 27th international Conference Extended Abstracts on Human Factors in Computing Systems Boston, MA, USA: CHI EA '09. ACM, New York, NY, 4811-4814. DOI=
<http://doi.acm.org/10.1145/1520340.1520747>
- Karam, M., Russo, F.A., & Fels, D.I. (2009, July 15). *Designing the model human cochlea: An ambient crossmodal audio-tactile display*. IEEE Transactions on Haptics. IEEE computer Society Digital Library. IEEE Computer Society,
<http://doi.ieeecomputersociety.org/10.1109/TOH.2009.32>
- Karam, M., Russo, F., Branje, C., Price, E., & Fels, D.I. (2008, May 28-20). Towards a model human cochlea: sensory substitution for crossmodal audio-tactile displays. In Proceedings of *Graphics interface 2008: vol. 322* (pp 267-274) Windsor, Ontario, Canada: GI Canadian Information Processing Society, Toronto, Ontario, Canada.
- Richman, W.L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*. 84(5). 754-775.
- Rose, D. H., & Meyer, A. (2002). *Teaching Every Student in the Digital Age*. Virginia: Association for Supervision and Curriculum Development, Alexandria.
- Shneiderman, B. (2000). Universal Usability. *Communications of the ACM* (43) 5: 84-91.
- Udo, J. P., & Fels, D. I. (2009). *The rogue poster children of universal design: Closed captioning and audio description*, Retrieved April 15, 2010 from Journal of Engineering Design:
<http://www.informaworld.com/smpp/content~db=all~content=a916629020>

About the Authors



Maria Karam

Karam's research focus is on human-computer interactions that are off-the-desktop, using alternative interaction techniques. Music is the main research area, ranging from gesture input to sensory substitution. Karam is also founder and director of the Coffee Lab, a public usability space in downtown Toronto.



Carmen Branje

Branje is a researcher at The Centre for Learning Technology at Ryerson University since 2003. His research interests are human factors engineering, human-computer interaction, cross modal translation, and multimodal interfaces. He will begin a PhD in Human Factors Engineering at the University of Toronto in the fall of 2009.



John-Patrick Udo

Udo has a Master's in Communication and Culture at Ryerson University and is currently a research fellow in the field of assistive and adaptive technology at the CLT. He is also an expert in universal design principles as applied to multimedia and entertainment technology.



Frank Russo

Russo, a cognitive scientist, musician, and armchair engineer, is also Assistant Professor of Psychology and Director of the Science of Music, Auditory Research, and Technology (SMART) lab. Russo sits on the editorial board of *Psychomusicology: Music, Mind, and Brain* and on the board of directors of the Canadian Acoustical Association and the Society for Music Perception, and Cognition.



Deborah I. Fels

Fels is a professor at Ryerson University and the head of the Centre for Learning Technologies (CLT). An expert in the field of assistive and adaptive technology, Fels and her research staff and students have been exploring accessible media, specifically enhanced and emotive captioning, and descriptive video since 1995.